

EXHIBIT C

Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211
Date: Wednesday, October 2, 2024 at 5:00:07 PM Pacific Daylight Time
From: Jordan Connors
To: Arbatman, Leeza, Craig Smyser, nyclassactions_microsoft_ohs@orrick.com, Alter-AI-TT@simplelists.susmangodfrey.com, MicrosoftNYClassActionFDBR@faegredrinker.com
CC: Dupree, Andrea, Geman, Rachel, Dozier, Wesley, Stoler, Reilly T., ssholder@cdas.com, ccole@cdas.com, OpenAICopyright, Lanham, John R., Christopher S. Sun, Andrew S. Bruns, Wood, Emily Claire, KVPOAI@keker.com, OpenAlcopyrightlitigation.lwteam@lw.com
Attachments: image003.png, image004.png, image005.png

Counsel, thank you for conferring with us on Monday regarding some of OpenAI's responses to our third and fourth sets of document requests. Please see below our takeaways from the call. We encourage you to correct anything that we have characterized inaccurately or that might need clarification.

We look forward to prompt production of the documents you agreed to search for and produce that are responsive to **RFP Nos. 36, 44, 46, 47, 48, 49, 57, 58, 66.**

RFP 29: You noted that this conversation is ongoing and has proceeded in a separate thread. You noted the agreement Leeza referenced below on September 28, 2024, and you agreed that, for any of the responsive "data access agreements and term sheets," by October 11 either a third party will file a motion to intervene or for a protective order or else OpenAI will produce the agreement or term sheet. Please note that Plaintiffs sent an email on this subject on September 25 and OpenAI has still not responded. Please provide a response to our proposal as soon as possible.

RFP 34-35: We noted that, given that OpenAI has not confirmed that the book-related datasets it has produced are Books1 and Books2, the spoliation issue, and therefore this request, is still live. OpenAI agreed to consider whether it would produce redacted retention agreements showing the date and scope of retention, and redacting any privileged material. Please let us know your position by Tuesday, October 8. As I stated on the call, I will let you know in advance if and when we decide to move to compel on this issue.

RFP 37: You indicated in your email that OpenAI would be willing to produce the requested documents if plaintiffs made a reciprocal production. We asked you on the call to explain what you meant by that, and to point to any requests

that relate to this issue or to any particular documents or things that we have refused to produce related to this issue. Please let us know by Tuesday, October 8.

RFP 45: We asked OpenAI to shed light on the volume of responsive documents to this request or to articulate any particular burden associated with producing them. We also noted that OpenAI's proposed limitation, depending on how it is applied to OpenAI's search and collection, might exclude relevant documents that speak generally to the conversion of OpenAI to a for-profit entity, but do not mention a GPT service by name. OpenAI explained that our formulation might include documents related to tax treatment that have no relevance to this dispute. OpenAI asked us to narrow the request.

We propose the following language to narrow the request, in response to the concern you expressed during our meet and confer:

All Documents Concerning or Relating to Your potential or actual conversion to a for-profit entity, which also discuss or reference (1) a GPT service or (2) the advantages, disadvantages, benefits, strategy, implications, or business projections associated with that potential or actual conversion. Additionally, this request excludes Documents directed solely to the tax implications of Your conversion to a for-profit entity.

RFP 46: You offered in an email below to conduct a reasonable search for documents "discussing whether and to what extent the relevant models reproduce, distribute, display, or create derivative works of copyrighted works." We asked whether your formulation of the request would include documents that discuss whether the use of copyrighted materials to train OpenAI LLMs complies with copyright law. You confirmed that the search you would conduct would include those documents as well. We look forward to receiving this production promptly.

RFP 51: You indicated that OpenAI is "still working on" a position with respect to RFP No. 51. You were unable to give us a date when we could expect to hear OpenAI's position, but you noted that you are aware that the parties are working against the backdrop of the Court's deadline to file an agenda for the

October 30 conference.

RFP 52: OpenAI asked us before the call to explain what we mean by outputs related to the Plaintiff works. We explained, as we had done on a previous meet and confer and in the email strong below, that we would like OpenAI to run author and class work names (e.g., as set forth by the “class work” terms in the parties’ custodial search term proposal) against retained LLM prompts and outputs. You responded that you are still investigating the technical feasibility of this request. Please let us know your position by Tuesday October 8.

RFP 54: You indicated that OpenAI believed this request pertained to a legal issue and that any responsive document about changing the terms and conditions would be privileged. We elaborated on the relevance of OpenAI’s terms and conditions (“ToC”) and explained that it appeared OpenAI actively changed the ToC to add an additional way to gather Copyrighted works for use in training data, which is relevant to Plaintiffs’ claims. We explained that because there appear to be business reasons behind the change, it is not clear that the requested documents would be privileged. You said that you would consider this and let us know whether you would search for and produce documents in response to this request. Please let us know by October 6.

RFP 90-94: You indicated that OpenAI would prefer not to produce any documents in response to these requests and asserted that only OpenAI OpCo is relevant to this lawsuit. You indicated that OpenAI would be willing to submit a declaration to that effect in lieu of producing documents. We explained that a declaration is not acceptable to Plaintiffs and that OpenAI should satisfy its discovery obligations with respect to these relevant requests, and we noted that, in a prior meet and confer, OpenAI’s counsel indicated that it would speak to the client about what documents it would be willing to produce in response to these requests.

As far as the relevance of these requests, in brief, they seek basic discovery concerning Plaintiffs’ claims against seven related entities that are all defendants in this lawsuit. The operative complaint identifies these entities and alleges facts as to each of them, which, if established, support claims for copyright infringement. Dkt No. 47, ¶¶ 49-58, 412-429. Many of those

allegations regard control and management of OpenAI or of OpenAI OpCo LLC by these other OpenAI Entities, together with knowledge of the infringing activity and financial interests in that same activity. See, e.g., Id. at 420, 421, 428. RFP Nos. 90-94 are targeted at documents and things that will support those allegations, such as transfers of assets and liabilities between and among these OpenAI entities (RFP Nos. 90 and 91), operational documents establishing a control structure among them, as alleged in the Complaint (RFP No. 94), and financial statements establishing the financial interests of each of these entities in the infringement alleged in the Complaint (RFP No. 93).

We assert these requests seek basic, highly relevant discovery that the Court will order OpenAI to produce if you force us to file a motion to compel. If you have evidence that any of these entities have been dissolved and never had control, management, ownership, or financial interests in the entities that have used or trained the LLMs at issue, please send us that evidence without further delay.

Kind regards,
Jordan

Jordan W Connors
Partner | SUSMAN GODFREY LLP
(O) 206-516-3814 | (C) 646-413-0743 | JConnors@SusmanGodfrey.com

From: Arbatman, Leeza (via Alter-AI-TT list) <Alter-AI-TT@simplelists.susmangodfrey.com>
Date: Saturday, September 28, 2024 at 2:49 PM
To: Craig Smyser <CSmyser@susmangodfrey.com>,
nyclassactions_microsoft_ohs@orrick.com <nyclassactions_microsoft_ohs@orrick.com>,
Alter-AI-TT@simplelists.susmangodfrey.com <Alter-AI-TT@simplelists.susmangodfrey.com>, MicrosoftNYClassActionFDBR@faegredrinker.com
<MicrosoftNYClassActionFDBR@faegredrinker.com>
Cc: Dupree, Andrea <ADupree@mofo.com>, Geman, Rachel <rgeman@lchb.com>,
Dozier, Wesley <wdozier@lchb.com>, Stoler, Reilly T. <rstoler@lchb.com>,
ssholder@cdas.com <ssholder@cdas.com>, ccole@cdas.com <ccole@cdas.com>,
OpenAICopyright <OpenAICopyright@mofo.com>, Lanham, John R.
<JLanham@mofo.com>, Christopher S. Sun <CSun@keker.com>, Andrew S. Bruns
<ABruns@keker.com>, Wood, Emily Claire <EWood@mofo.com>, KVPOAI@keker.com
<KVPOAI@keker.com>, OpenAlcopyrightlitigation.lwteam@lw.com
<openaicopyrightlitigation.lwteam@lw.com>

Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Hi Craig,

As promised, here is a summary of our positions on pending RFPs in Plaintiffs' third set. We look forward to discussing these proposals with you during the meet and confer on Monday (9/30).

Data access agreements. In **RFP 29**, Plaintiffs seek all documents "sufficient to show any deals, contracts, [and] agreements . . . concerning the licensing or purchase of data . . . to train ChatGPT." As OpenAI has previously stated, it has already produced certain data access agreements and term sheets. It is in the process of providing necessary notice to relevant third parties and will produce additional agreements in the next few weeks. In its notice letters, OpenAI is informing the third parties that if they object to OpenAI's production of agreements or communications related to training data, they should file a motion to intervene or for a protective order by October 11, 2024 to protect their interests. We understand that the discussion regarding this RFP is happening in a separate thread, but wanted to reiterate our position here for clarity.

Retention agreements. In **RFPs 34 and 35**, Plaintiffs seek retention agreements between OpenAI and any law firms (without any temporal limitation), as well as retention agreements with MoFo and Latham between 2016 and 2022. Plaintiffs assert that OpenAI's alleged spoliation of evidence (i.e., the deletion of certain book-related datasets due to nonuse) makes these retention agreements relevant and that the scope of retention is not privileged. While OpenAI maintains that Plaintiffs' spoliation argument has always been unfounded, that argument now holds no water at all, given OpenAI's production of the book-related datasets. OpenAI believes that this production renders the requested retention agreements irrelevant. We therefore propose tabling this issue while Plaintiffs review these datasets.

Communications with third parties. In **RFP 36**, Plaintiffs seek OpenAI's communications with third parties between 2016-2022 relating to the use of copyrighted materials to train OpenAI's artificial intelligence models. As stated in OpenAI's correspondence regarding RFPs 7-9, OpenAI has agreed to conduct a reasonable search for non-privileged documents in our possession, custody, or control discussing whether and to what extent the relevant models reproduce, distribute, display, or create derivative works of copyrighted works.

Documents related to USPTO Comment on IP for AI innovation. In **RFP 37**, Plaintiffs seek all documents relating to the preparation of the comment OpenAI submitted in response to the USPTO's "request for comments on intellectual property protection for [AI] innovation," including documents provided to those who drafted the comment. OpenAI will conduct a reasonable search for non-privileged documents in its possession, custody, or control that were relied upon in preparing the comment, provided that Plaintiffs agree to a reciprocal production.

Conversion to for-profit entity. In **RFP 45**, Plaintiffs seek all documents concerning OpenAI's "potential or actual conversion to a for-profit entity." This request is overly burdensome and seeks irrelevant information. But as stated in the September discovery conference, in the interest of compromise, OpenAI will produce documents about the decision to create a for-profit entity that also discuss or address a GPT service.

Copyrighted material. In **RFP 46**, Plaintiffs seek all documents concerning whether OpenAI's use of copyrighted material in its LLMs complies with copyright law in the U.S. and elsewhere. This request is

overly burdensome and seeks privileged and/or irrelevant information. Nevertheless, in the interest of compromise, OpenAI will conduct a reasonable search for non-privileged documents in its possession, custody, or control discussing whether and to what extent the relevant models reproduce, distribute, display, or create derivative works of copyrighted works.

Microsoft's investment into OpenAI. In **RFPs 47 and 48**, Plaintiffs seek all documents relating to Microsoft's investments into OpenAI and the inclusion of OpenAI's LLMs, including ChatGPT, in Microsoft's products, as well as "profits or revenues from such products and the distribution thereof" among OpenAI and Microsoft. OpenAI believes these requests are overbroad, including because they seek information about models not at issue in this litigation. Nevertheless, in the interest of compromise, OpenAI will conduct a reasonable search for non-privileged documents in its possession, custody, or control discussing (i) products, services, or "in-kind" investments provided by Microsoft specifically for OpenAI's development, deployment, commercialization, and training of the relevant models, including documents discussing the importance of the same; and (ii) integrating, or the possibility of integrating, ChatGPT or the relevant models into Microsoft products.

Commercialization of ChatGPT. In **RFP 49**, Plaintiffs seek all documents relating "to the use of inclusion of ChatGPT or [OpenAI's LLMs] in commercial products, including but not limited to documents concerning or relating to the ideation, development, or commercialization timeframe." In their 9/20 correspondence, Plaintiffs asserted that OpenAI never provided its position on this RFP. We are confused by that assertion. OpenAI made an offer to Plaintiffs in an email one day earlier (grouped with its offer for RFP 44). In fact, you copied and pasted OpenAI's position and seemingly accepted its offer in your correspondence below.

Balance sheets and cash flow. In **RFP 51**, Plaintiffs seek the "balance sheet, income statement, and cash flow statement, on a monthly basis," for each OpenAI entity. As previously stated, OpenAI OpCo is the operating entity and thus the relevant entity in this litigation. The other entities do not play any role in the development of the models at issue. We are still working on formulating a proposal for this request and will follow up separately.

Outputs related to Plaintiffs' works. In **RFP 52**, Plaintiffs seek all documents sufficient "to show the output of any [LLMs] created by [OpenAI] related to Plaintiffs' works." OpenAI is still investigating the technical feasibility and burden of searching for responsive documents. OpenAI requests that during the meet and confer on Monday, Plaintiffs clarify what they mean by outputs "related to" Plaintiffs' works, as the answer will affect our investigation.

ChatGPT terms of use. In **RFP 54**, Plaintiffs seek all documents concerning OpenAI's decision to include, modify, or remove specific language from ChatGPT's terms and conditions asserting that OpenAI should log the privileged documents. That is not a solution. Locating a non-negligible quantity of documents we anticipate to be fully privileged and logging them will require expending significant time and resources. That will do nothing but burden OpenAI. Moreover, Plaintiffs' theory of relevance remains unclear. OpenAI is willing to meet and confer on this issue to better understand how all documents concerning or relating to the aforementioned language in the terms and conditions is relevant.

Future impact of ChatGPT. In **RFP 57**, Plaintiffs seek all documents relating "to the impact or potential future impact of [OpenAI's LLMs], ChatGPT, or other artificial intelligence on the [] labor, art, or writing [markets]." OpenAI will conduct a reasonable search for relevant, non-privileged documents in

its possession, custody, and control discussing the impact or potential impact of LLMs on the market for labor, art, or writing.

Creator compensation. In **RFP 58**, Plaintiffs seek all documents relating to OpenAI's plans to allow creators to receive compensation or determine how and whether their works will be used by OpenAI, including documents related to Sam Altman's Congressional testimony that "creators deserve control over how their creations are used." OpenAI will conduct a reasonable search for relevant, non-privileged documents in its possession, custody, and control discussing (i) the phrase "creators deserve control over how their creations are used" in Sam Altman's congressional testimony and (ii) as a general matter, whether and how to allow authors, creators, or copyright holders to receive compensation to the extent their copyrighted works were included in the data on which the relevant models were trained.

Leeza Arbatman

Associate

LArbatman@mofo.com

T +1 (925) 766-6920

**MORRISON
FOERSTER**

From: Craig Smyser <Csmyser@susmangodfrey.com>

Sent: Wednesday, September 25, 2024 7:28 PM

To: Arbatman, Leeza <LArbatman@mofo.com>; Wood, Emily Claire <EWood@mofo.com>; nyclassactions_microsoft_ohs@orrick.com; MicrosoftNYClassActionFDBR@faegredrinker.com; OpenAICopyrightlitigation.lwteam@lw.com; KVPOAI@keker.com

Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley <wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; sshoulder@cdas.com; ccole@cdas.com; Alter-AI-TT@simplelists.susmangodfrey.com; Ranieri, Vera <VRanieri@mofo.com>; OpenAICopyright <OpenAICopyright@mofo.com>

Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

Importance: High

External Email

We won't reiterate our prior correspondence, however we can be available at 11 AM PT on 9/30. Please send through an invitation.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001

(212) 336-2023 (office) | (713) 503 6252 (cell)

csmyser@susmangodfrey.com | www.susmangodfrey.com

From: "Arbatman, Leeza" <LArbatman@mofo.com>
Date: Wednesday, September 25, 2024 at 8:24 PM
To: Craig Smyser <CSmyser@susmangodfrey.com>, "Wood, Emily Claire" <EWood@mofo.com>, "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>, "MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, "OpenAlcopyrightlitigation.lwteam@lw.com" <openaicopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com" <ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, Alter-TT <Alter-AI-TT@simplelists.susmangodfrey.com>, "Ranieri, Vera" <VRanieri@mofo.com>, OpenAICopyright <OpenAlCopyright@mofo.com>
Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Hi Craig,

Thanks for letting us know Plaintiffs' positions. We don't feel that we're at an impasse and are diligently working on preparing additional proposals. We think there are still some open items that would be helpful to discuss live. Are you available on Monday (9/30) for a conference? We will be sure to send our proposals in writing before we meet. Let us know.

Thanks,
Leeza

Leeza Arbatman
Associate
LArbatman@mofo.com
T +1 (925) 766-6920

**MORRISON
FOERSTER**

From: Craig Smyser <CSmyser@susmangodfrey.com>
Sent: Friday, September 20, 2024 4:56 PM
To: Wood, Emily Claire <EWood@mofo.com>; Ranieri, Vera <VRanieri@mofo.com>; nyclassactions_microsoft_ohs@orrick.com; MicrosoftNYClassActionFDBR@faegredrinker.com; OpenAICopyright <OpenAlCopyright@mofo.com>; OpenAlcopyrightlitigation.lwteam@lw.com; KVPOAI@keker.com
Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley <wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; ssholder@cdas.com; ccole@cdas.com; Alter-AI-TT@simplelists.susmangodfrey.com

Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

External Email

Emily and Leeza,

I have pasted below the portions of your response sent 9/19 2:03 ET on the thread discussing Plaintiffs' Fourth RFPs but which concern Plaintiffs' Third RFPs (the email is also attached). Please keep the threads separate (including any forthcoming response on "RFPs 50, 52, 59-63, and 65-70" promised in your 9/19 email). It is important to preserve the timeline of the meet and confer process and the parties' exchanges for the two separate set of RFPs for the Court and because, as shown below, it appears counsel for OpenAI are not communicating regarding meet and confers and the parties' prior correspondence and commitments.

At the outset, we note that the vast majority of questions which OpenAI poses in its response were already addressed on meet and confers or in prior correspondence. Your response appears to ignore all previous correspondence and meeting and conferring on these issues. Particularly in light of the Court's instructions at the most recent status conference, it is unacceptable for OpenAI to stall discovery by asking for explanations already provided and failing to provide responses to questions which it previously promised to answer. We intend to note as much in our forthcoming status conference letter. As stated previously, OpenAI needs to respond to the issues laid out in my 8/27 and 7/29 emails.

The following is a brief response to each of the points made in your 9/19 email dealing with the third RFPs:

34 and 35. You have not responded to my email of 7/29 or the explanation provided on the meet and confer of 7/26 regarding spoliation. From my 7/29 email:

"Regarding Requests 34 and 35, you had asked us to send case law on the discoverability of retention agreements in light of the deletion of Books1 and Books2. *See, e.g., Linet Americas, Inc. v. Hill-Rom Holdings, Inc.*, 2023 WL 9119836, at *1-*3 (N.D.Ill., 2023) (deletion of relevant documents was a sufficient factual basis to permit discovery of the "retention letter and a log containing the dates, participants in, and subject matter of communications between Linet and its counsel related to potential claims against Hillrom from December 2019 to April 2020"). The scope of that retention is not privileged. *Id.; Berall v. Teleflex Medical Incorporated*, 2021 WL 5989936, at *2 (S.D.N.Y., 2021)."

The parties are at an impasse and we will consider moving the Court to compel however please inform us by next Wednesday if OpenAI has made any progress here.

36: Thank you, the parties are at an impasse and we will consider moving the Court to compel, if you wish to identify specific categories of documents which OpenAI has already committed to

produce on this issue which you believe satisfies your obligations, please do so by next Wednesday.

37: You ask “Is there a substantive issue covered by the comment that you're seeking more information about that is relevant to the issues in this case, and which has not been otherwise provided via other RFPs” This question again ignores the parties' previous meet and confer—in which we answered this question--and my 7/29 email:

“Regarding request 37 (documents regarding the preparation of OpenAI's comment to the USPTO), your position was that OpenAI would be willing to search for and produce in response to this request if Plaintiffs' did the same regarding certain RFPs directed to their statements regarding AI. Please let me know if that accurately portrays OpenAI's position, and, if so, which specific RFPs to Plaintiffs are being referenced.”

Please respond to the question in my 7/29 email, which has been awaiting answer for nearly two months. In the meantime, we consider the parties at an impasse and will consider moving the court, however please inform us by next Wednesday if OpenAI has made any progress here.

38: thank you for confirming that OpenAI is only willing to produce the agreements but not the communications themselves. We will review such documents and consider if it satisfies this RFP.

39-40: We await your production of “documents in our possession, custody, and control discussing (i) how and why OpenAI curates and uses different types of data to train OpenAI's relevant models; and (ii) the relative value of types of data for use in the relevant models.” We will review such documents and consider if that satisfies these RFPs.

41-43: We await your production of “documents in our possession, custody, and control discussing or describing (i) the techniques, technologies, or procedures OpenAI employs to prevent the relevant models from regurgitating training data; (ii) the reasons for implementing those techniques, technologies, or procedures; (iii) OpenAI's creation of human-generated text for the purpose of acting as training data for the relevant models; and (iv) the changes in output quality, if any, of the relevant models, based on the use of synthetic data to train these models as opposed to human-generated text.” We will review such documents and consider if that satisfies these RFPs.

44 and 49: We await your production of “documents in our possession, custody, and control discussing or describing (i) how ChatGPT and relevant models can be or are used for OpenAI commercial uses and applications; and (ii) OpenAI's decision to integrate ChatGPT or relevant models in OpenAI's commercial products.” We will review such documents and consider if that satisfies these RFPs.

45: you ask “*Please explain why this data is relevant to Plaintiffs' claims or damages.*” This question again ignores the parties' previous meet and confer and my 7/29 email:

“Regarding request 45 (documents relating to conversion to for-profit)*, you were going to confirm with your client that their position remains the same as in the *New York Times* case. We will review the briefing from the *New York Times* case on this issue and let[] you know our position on the parties mutually agreeing to be bound by the Court's

eventual decision on that motion.”

Please respond to the question in my 7/29 email, which has been awaiting answer for nearly two months. We further note that we have set forth our position on this issue and our willingness to be consider OpenAI’s approach to resolving its similar dispute with the New York Times via our participation in today’s joint meet and confer with the Times.

46: The parties are at an impasse. Prior to moving the Court, we note again that you have failed to respond to my 7/29 email setting forth the relevant caselaw which OpenAI specifically requested. If you have such caselaw, please provide it by next Wednesday:

“Regarding request 46 (documents regarding whether your use of copyrighted material in LLMs complies with copyright law in the United States or elsewhere), you will look at the current search terms to determine your view on whether the relevant non-privileged documents responsive to this request would already be captured and, to the extent they were not, propose an alternative scope as to the portion of the request dealing with compliance with United States Copyright law. You had asked if we had case law supporting the relevance of whether OpenAI’s use of copyrighted material complies with non-US Copyright laws. Please see *Bluewater Music Servs. Corp. v. Spotify USA Inc.*, No. 3:17-CV-01051, 2018 WL 10152456, at *1, *8-9 (M.D. Tenn. July 10, 2018) (ordering production of documents showing that Defendant was “able to appropriately pay royalties for the same musical compositions to music rights organizations in other countries” over Defendant’s objection that compositions by entities “operating under foreign laws, would [not] put Defendant on notice of its alleged non-compliance with U.S. copyright law . . .”).”

47: We await the production of “(i) executed agreements with Microsoft regarding their investment in OpenAI or OpenAI’s development of LLMs; and (ii) agreements between Microsoft and OpenAI regarding the integration of ChatGPT into Microsoft products,” however please respond to my 7/29 email on this issue:

“Regarding request 47 (documents regarding products, services, or “in-kind” investments provided to OpenAI by Microsoft and the importance of the same), you had asked us to propose a narrowed scope. Plaintiffs are not interested in whether OpenAI is using Microsoft Word as its default word processing software or OpenAI’s general usage of Microsoft’s products. To the extent OpenAI had interpreted the request as including such documents, we agree to narrow the scope. Plaintiffs rather seek the documents regarding the products, services, or “in-kind” investments provided by Microsoft specifically for OpenAI’s development, deployment, commercialization, and training of large language models, as well as the importance of the same. Please let us know if OpenAI will produce these documents.”

Your email fails to acknowledge this prior exchange. We thus understand the parties are at an impasse and will consider moving the court however please inform us by next Wednesday if OpenAI has made any progress here.

48: We await the production of “(i) executed agreements with Microsoft regarding their investment in OpenAI or OpenAI’s development of LLMs; and (ii) agreements between Microsoft and OpenAI regarding the integration of ChatGPT into Microsoft products,”

however your response is unacceptable given that Ms. Wood's email of 8/26 acknowledge Plaintiffs' modified scope (from an email sent 8/20) of:

"(1) documents reflecting the decision to integrate ChatGPT or OAI's Large Language Models in Microsoft's Products, (2) documents reflecting the evaluation of Microsoft's Azure platform for serving / deploying ChatGPT or OAI's Large Language Models as part of Microsoft's Products and (3) documents showing the profits or revenues from the integration of ChatGPT or OAI's Large Language Models in Microsoft's Products. We consider this proposal to be Plaintiff's current position as to the documents it believes are necessary to consider the RFP satisfied."

Ms. Wood's 8/26 email asked Plaintiffs to clarify which products were covered by the scope to which we promptly responded on 8/27 email:

"With respect to request 48, OpenAI is best suited to identify which Microsoft products incorporate its LLMs and/or ChatGPT. We expect OpenAI to produce for those products which it knows use OpenAI's LLMs and/or ChatGPT. We would at least expect it to include those products for which public reporting states that OpenAI's LLM's have been used, e.g., Bing (<https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/>), Edge (<https://www.microsoft.com/en-us/edge/features/ai?form=MA13FJ>), GitHub Copilot (<https://github.com/features/copilot>), Microsoft 365 Copilot (<https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/>), Power Platform (<https://techcommunity.microsoft.com/t5/apps-on-azure-blog/build-next-gen-apps-with-openai-and-microsoft-power-platform/ba-p/3742289>), Dynamics 365 CoPilot (<https://techcrunch.com/2023/03/06/microsoft-dynamics-copilot/>), and Azure OpenAI Service (https://learn.microsoft.com/en-us/azure/ai-services/openai/overview?WT.mc_id=academic-105485-koreyst). Please provide your response by next Tuesday."

Your email fails to acknowledge or deal with this prior exchange. We thus understand the parties are at an impasse and will consider moving the court however please inform us by next Wednesday if OpenAI has made any progress here.

49: your email fails to mention RFP 49, this was covered in my 8/20 email summarizing the parties' 8/16 meet and confer in which we stated:

"we propose that OpenAI search for and produce (1) documents showing the decision to integrate ChatGPT or OAI's Large Language Models in commercial products, (2) documents showing the development of the particularly commercial aspect of those products (as distinct from their general development as LLMs), and (3) documents showing the timeline / speed of development. Please let us know your position on these three categories."

We again requested an answer on 8/27. It is unacceptable that OpenAI has failed to provide its position. We consider the parties at an impasse and will consider moving the Court however please inform us by next Wednesday if OpenAI has made any progress here.

51: This was covered in the parties' 8/16 meet and confer in which Plaintiffs explained the relevance in question (I note that while we disagree with the summaries, Ms. Wood's email of

8/26 and my email of 8/20 clearly confirms that the parties have already discussed this issue) and counsel promised to ask OpenAI what it would be willing to produce on this. Your response suggests counsel has not done that. It is unacceptable that OpenAI has failed to do so. We consider the parties at an impasse and will consider moving the Court however please inform us by next Wednesday if OpenAI has made any progress here.

52: your email fails to mention RFP 52, which was covered in my 8/20 email summarizing the parties' 8/16 meet and confer in which I stated:

"With respect to request 52, we suggest that, as an initial step, you search the output data for the named plaintiff's works and/or their names. Depending on the volume which results, OpenAI could either produce this data or make it available for inspection as part of the training data / model inspection. Please let us know your position on this."

We again requested an answer on 8/27. While we note the forthcoming response mentioned in your 9/19 email, it is unacceptable that OpenAI has still failed to provide its position. We consider the parties at an impasse and will consider moving the Court however please inform us by next Wednesday if OpenAI has made any progress here.

53: We await production of "all final versions of the terms and conditions of use for ChatGPT within our possession, custody, or control." We will review such documents and consider if that satisfies this RFP.

54: please respond to my 7/29 email on this issue:

"Regarding request 54 (documents concerning the addition of certain language to OpenAI's terms and conditions for ChatGPT), you had asked us to identify the relevance of this material and consider a revised scope. As to a revised scope, we believe the request is appropriate as written.

As to relevance, OpenAI's terms and conditions state that "By uploading any User Content you hereby grant and will grant OpenAI and its affiliated companies a nonexclusive, worldwide, royalty free, fully paid up, transferable, sublicensable, perpetual, irrevocable license to copy, display, upload, perform, distribute, store, modify and otherwise use your User Content **for any OpenAI-related purpose** in any form, medium or technology now known or later developed." As stated on the call, we believe that OpenAI could claim to be authorized to, e.g., train its models on such user submitted content. We further note that OpenAI claims exactly that on its website:

"Will OpenAI use files uploaded to train its models?

The answer depends on the service you are using. As explained in [this article](#), we may use content submitted to ChatGPT, DALL·E, and our other services for individuals to improve model performance. Content may include files that are uploaded. Please refer to this [article](#) to understand how content may be used to improve model performance and the choices that users have."

https://help.openai.com/en/articles/8555545-file-uploads-faq#h_3137baa34c

That is relevant to this case as it may be an additional source of Plaintiffs' copyrighted material used in training OpenAI's models."

Your email fails to acknowledge this prior exchange. We further note that whether information is privileged is not a basis for refusing to search for the material in question, but rather a basis for logging it and producing a log and/or redacted non-privileged portions of documents. We understand the parties are at an impasse and will consider moving the court however please inform us by next Wednesday if OpenAI has made any progress here.

55: We await production of documents “discussing how and why OpenAI curates and uses different types of data (including data, if any, from Library Genesis, LibGen, Internet Archive, Z-Library, Common Crawl, WebText, Project Gutenberg, Anna’s Archive, Open Library, Reddit, or DuXiu) to train the relevant models.” We will review such documents and consider if that satisfies this RFP.

56: We await production of documents referencing “Fair Learning.” We will review such documents and consider if that satisfies this RFP..

57: We await production of documents “in our possession, custody, and control discussing the impact or potential impact of LLMs on the market for books.” However, in our email of 8/27, we stated

“With respect to request 57, we believe the request is fine as written given the large number of public statements by OpenAI leadership regarding these exact issues, e.g., <https://www.hbs.edu/bigs/sam-altman-harvard-business-school>, <https://openai.com/index/gpts-are-gpts/>, <https://arxiv.org/abs/2303.10130>, <https://fortune.com/2023/10/18/openai-sam-altman-misjudged-ai-creativity-chatgpt/>, <https://www.sfchronicle.com/bayarea/article/openai-funds-research-impact-universal-basic-19589117.php>, <https://www.techtimes.com/articles/304379/20240507/openai-ceo-sam-altman-admits-ai-eliminate-current-jobs.htm>, <https://openai.com/index/economic-impacts/>, <https://openai.com/global-affairs/testimony-of-sam-altman-before-the-us-senate/>, and <https://openai.com/index/language-model-safety-and-misuse/>. Please let us know if you will search for and produce documents responsive to this RFP by next Tuesday.”

Similarly, we note that OpenAI itself has taken the position that the benefits and costs of its models are relevant to the fair use factors in its briefing in the NYT case. As such it is inappropriate for OpenAI to withhold documents relating to AI’s impact or potential impact of LLMs on the labor or art markets. We understand the parties are at an impasse and will consider moving the court however please inform us by next Wednesday if OpenAI has made any progress here.

58: We await production of “documents in our possession, custody, and control discussing or describing, as a general matter, whether and how to address preferences by authors, creators, or rightsholders regarding their data that may be used in training the relevant models, including opt-outs related to data.” However, in our email of 8/27 email we stated:

“For RFP 58, you requested that we put categories in writing that you can pose to your client. The categories are: (1) drafts or discussion of the phrase “creators deserve control

over how their creations are used" in Sam Altman's congressional testimony and (2) your internal discussions regarding whether and how to compensate authors or rightsholders and/or enable opt-out of the use of their data in training."

It is unacceptable that OpenAI has failed to provide its position in response to this previous meet and confer. We consider the parties at an impasse and will consider moving the Court however please inform us by next Wednesday if OpenAI has made any progress here.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-2023 (office) | (713) 503 6252 (cell)
csmysyer@susmangodfrey.com | www.susmangodfrey.com

**Portions of your response sent 9/19 2:03 ET on the thread discussing Plaintiffs' Fourth RFPs but which concern Plaintiffs' Third RFPs

Please find below (i) our responses to RFPs 33-49, 51, and 53-58; and (ii) a summary of our meet and confer from Monday, September 9. We will follow up regarding RFPs 50, 52, 59-63, and 65-70.

Responses to RFPs 33-49, 51, and 53-58

Retention agreements. In **RFPs 34 and 35**, you ask for retention agreements between OpenAI and any law firms over any time period, as well as retention agreements with MoFo and Latham between 2016 and 2022. We believe these requests are overbroad, including because (i) retention agreements from any time period will relate to matters not at issue in this litigation, and (ii) agreements between MoFo/Latham and OpenAI over an 8-year period starting before this litigation was initiated will also relate to matters not at issue in this litigation. We believe these requests also seek irrelevant information, as OpenAI's engagement with different law firms over the course of its existence has no bearing on the copyright issues raised here. Moreover, to the extent you seek information about OpenAI's relationship with its lawyers and the reasons OpenAI retained them, such information is privileged. Finally, Plaintiffs have not established a need for this information, especially in light of OpenAI's productions or anticipated productions. We will therefore not produce any documents in response. If you have some theory as to why these documents are relevant and/or you are entitled to them, please identify what that reason is.

Communications with third parties. In **RFP 36**, you ask for OpenAI's communications with third parties between 2016-2022 relating to the use of copyrighted materials to train OpenAI's artificial intelligence models. We have conducted a reasonable search and, as we have continued to represent, there is no centralized repository for third-party complaints, nor is there an employee primarily responsible for fielding these issues. We believe that a request that counsel search across the entire company is unduly

burdensome and not proportional to the needs of the case. We will therefore not produce any additional documents in response to this request, beyond those we have previously agreed to produce.

Documents related to USPTO Comment on IP for AI innovation. In **RFP 37**, you ask for all documents relating to the preparation of the comment OpenAI submitted in response to the USPTO request for comments on intellectual property protection for AI innovation, including documents provided to those who drafted the comment. We believe this request seeks privileged information to the extent it seeks attorney work product, or documents provided to OpenAI employees by their legal counsel. Moreover, the burden of collecting and reviewing possibly responsive documents vastly outweighs any relevance, which OpenAI disputes. Nevertheless, in the interest of compromise, we are willing to consider your request if you narrow the scope of the documents that you seek and establish their relevance. *Is there a substantive issue covered by the comment that you're seeking more information about that is relevant to the issues in this case, and which has not been otherwise provided via other RFPs?*

Data access agreements and term sheets. In **RFP 38**, you ask for all communications between OpenAI and anyone else regarding the use of that's party's text or audio in connection with its LLMs. As we have stated previously, we have already produced certain data access agreements and term sheets. We are in the process of providing notice to third parties prior to producing additional agreements. Accordingly, we will supplement our production once outstanding issues have been resolved.

Type and value of data. In **RFPs 39 and 40**, you ask for all documents concerning the types and values of types of data used to train OpenAI's LLMs. We believe these requests are overbroad, including because they seek information about models not at issue in this litigation. Nevertheless, in the interest of compromise, we will conduct a reasonable search for non-privileged documents in our possession, custody, and control discussing (i) how and why OpenAI curates and uses different types of data to train OpenAI's relevant models; and (ii) the relative value of types of data for use in the relevant models.

Training models. In **RFPs 41-43**, you ask for all documents concerning (i) modifying any parameters for tuning or restricting the output of LLMs or ChatGPT to avoid copyright infringement; (ii) creating human-generated text for use as training data for LLMs; and (iii) changes in output quality of LLMs based on the inclusion/exclusion of human-generated text in the training data for LLMs. We believe these requests are overbroad, including because they seek information about models not at issue in this litigation. Nevertheless, in the interest of compromise, we will conduct a reasonable search for relevant and non-privileged documents in our possession, custody, and control discussing or describing (i) the techniques, technologies, or procedures OpenAI employs to prevent the relevant models from regurgitating training data; (ii) the reasons for implementing those techniques, technologies, or procedures; (iii) OpenAI's creation of human-generated text for the purpose of acting as training data for the relevant models; and (iv) the changes in output quality, if any, of the relevant models, based on the use of synthetic data to train these models as opposed to human-generated text.

Commercialization of ChatGPT/LLMs. In **RFPs 44 and 49**, you ask for all documents

concerning the commercial uses and/or applications of ChatGPT and other LLM, including documents concerning a commercialization timeframe. We believe these requests are overbroad, including because they seek information about models not at issue in this litigation. Nevertheless, in the interest of compromise, we will conduct a reasonable search for relevant and non-privileged documents in our possession, custody, and control discussing or describing (i) how ChatGPT and relevant models can be or are used for OpenAI commercial uses and applications; and (ii) OpenAI's decision to integrate ChatGPT or relevant models in OpenAI's commercial products.

Conversion to for-profit entity. In **RFP 45**, you ask for all documents concerning OpenAI's potential or actual conversion to a for-profit entity. This request is overly burdensome and seeks irrelevant information. It is unclear how documents relating to the decision to adopt a particular corporate form are relevant to the issues in this litigation. The burden of seeking all such documents is significant. *Please explain why this data is relevant to Plaintiffs' claims or damages.*

Copyrighted material. In **RFP 46**, you ask for all documents concerning whether OpenAI's use of copyrighted material in its LLMs complies with copyright law in the U.S. and elsewhere. This request is overly burdensome and seeks privileged and/or irrelevant information. To the extent any such documents exist, OpenAI's legal team almost certainly drafted them, so they are privileged. And to the extent you seek all documents related to OpenAI's compliance with copyright laws in relation to its LLMs, it is unclear how this is relevant to the specific issues in this litigation. The burden of searching for all such documents is significant. We therefore refuse to produce documents in response.

Microsoft's investment into OpenAI. In **RFPs 47 and 48**, you ask for all documents relating to Microsoft's investments into OpenAI and the inclusion of OpenAI's LLMs, including ChatGPT, in Microsoft's products, as well as profits or revenues from such products and the distribution of the same between OpenAI and Microsoft. We believe these requests are overbroad, including because they seek information about models not at issue in this litigation. Nevertheless, in the interest of compromise, we will produce (i) executed agreements with Microsoft regarding their investment in OpenAI or OpenAI's development of LLMs; and (ii) agreements between Microsoft and OpenAI regarding the integration of ChatGPT into Microsoft products.

Balance sheets and cash flow. In **RFP 51**, you ask for the balance sheet, income statement, and cash flow statement, on a monthly basis, for each OpenAI entity. We believe this request is overbroad and seeks irrelevant information. It is unclear how such financial information relates to Plaintiffs' claims or defenses. Plaintiffs also seek such information from entities not at issue in this litigation. *What is the specific financial information that you seek and how is it relevant?*

ChatGPT terms of use. In **RFPs 53 and 54**, you ask for all documents showing each version of the terms and conditions of use for ChatGPT, as well as all documents concerning OpenAI's decision to include, modify, or remove language from ChatGPT's terms and conditions that represents that a user of ChatGPT owns all rights to any user content that it uploads to ChatGPT, and that, by uploading the user content, the user grants OpenAI full rights to copy, display, distribute etc. the user content for OpenAI-related

purposes in any form or medium. We will conduct a reasonable search of all final versions of the terms and conditions of use for ChatGPT within our possession, custody, or control. Regarding your request for documents relating to the decision to include the language discussed above, we believe this request seeks privileged information, as discussions surrounding whether and how to use such language involved the advice of counsel and attorney work product. We will thus not produce any documents in response.

Repositories of text used to train LLMs. In **RFP 55**, you ask for all documents relating to any version of Library Genesis, LibGen, Internet Archive, Z-Library, Common Crawl, WebText, Project Gutenberg, Anna's Archive, Open Library, Reddit, DuXiu, or any other repository of text used to train OpenAI's LLMs. We believe this request is overbroad, including because it seeks information about models not at issue in this litigation. This request also seeks privileged information to the extent it seeks communications regarding these text repositories between OpenAI and its legal counsel. Nevertheless, in the interest of compromise, we will conduct a reasonable search for non-privileged documents in our possession, custody, or control discussing how and why OpenAI curates and uses different types of data (including data, if any, from Library Genesis, LibGen, Internet Archive, Z-Library, Common Crawl, WebText, Project Gutenberg, Anna's Archive, Open Library, Reddit, or DuXiu) to train the relevant models.

"Fair Learning." In **RFP 56**, you ask for all documents relating to the academic paper "Fair Learning," by Mark Lemley and Bryan Casey. In the interest of compromise, we will conduct a reasonable search for non-privileged documents in our possession, custody, and control referencing this paper.

Future impact of ChatGPT. In **RFP 57**, you ask for all documents relating to the impact or potential impact of OpenAI's LLMs, ChatGPT, or other artificial intelligence on the labor, art, or writing markets. This request is overbroad and seeks irrelevant information. All documents relating to the impact of LLMs, ChatGPT, or artificial intelligence on the labor, art, or writing markets implicates information far beyond the scope of this litigation. Nevertheless, in the interest of compromise, we will conduct a reasonable search for relevant, non-privileged documents in our possession, custody, and control discussing the impact or potential impact of LLMs on the market for books, as that is the only market implicated by this lawsuit.

Creator compensation. In **RFP 58**, you ask for all documents relating to OpenAI's plans to allow creators to receive compensation or determine how and whether their works will be used by OpenAI, including documents related to Sam Altman's testimony that "creators deserve control over how their creations are used." We believe this request is overbroad to the extent it seeks *any conversations* with creators related to the subject of Sam Altman's testimony. Nevertheless, in the interest of compromise, we will conduct a reasonable search for relevant, non-privileged documents in our possession, custody, and control discussing or describing, as a general matter, whether and how to address preferences by authors, creators, or rightsholders regarding their data that may be used in training **the relevant models**, including opt-outs related to data.

From: "Wood, Emily Claire" <EWood@mofo.com>
Date: Monday, September 16, 2024 at 6:11 PM
To: Craig Smyser <CSmyser@susmangodfrey.com>, "Ranieri, Vera" <VRanieri@mofo.com>, "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>, "MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright <OpenAICopyright@mofo.com>, "OpenAICopyrightlitigation.lwteam@lw.com" <OpenAICopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com" <ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, Alter-TT <Alter-AI-TT@simplelists.susmangodfrey.com>
Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Craig,

We responded in another thread (attached), but to be clear, that response applies here as well. We are working on our responses to Plaintiffs' RFPs in light of the Court's guidance at the omnibus discovery hearing last Thursday. We will get back to you soon.

Emily

From: Craig Smyser <CSmyser@susmangodfrey.com>
Sent: Monday, September 16, 2024 5:22 PM
To: Wood, Emily Claire <EWood@mofo.com>; Ranieri, Vera <VRanieri@mofo.com>; nyclassactions_microsoft_ohs@orrick.com; MicrosoftNYClassActionFDBR@faegredrinker.com; OpenAICopyright <OpenAICopyright@mofo.com>; OpenAICopyrightlitigation.lwteam@lw.com; KVPOAI@keker.com
Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley <wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; ssholder@cdas.com; ccole@cdas.com; Alter-AI-TT@simplelists.susmangodfrey.com
Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

External Email

Counsel:

OpenAI served its responses and objections to Plaintiffs' Third RFPs on July 15th and took more than a month to find time to meet and confer with Plaintiffs regarding each issue

OpenAI raised in its responses. It is now September 16th—more than two months after the responses were served, 49 days since my 7/29 correspondence, and 20 days since my 8/27 correspondence--and we are still awaiting information, answers, and OpenAI's positions on nearly all of these requests.

These delays prevent expedient identification and resolution of discovery issues. We intend to inform the Court of this pattern of obstructive delays in our September 27, 2024 Joint Status letter.

Please provide OpenAI's response on the issues in my 8/27 and 7/29 correspondence by this Friday.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-2023 (office) | (713) 503 6252 (cell)
csmysyer@susmangodfrey.com | www.susmangodfrey.com

From: Alter-TT <Alter-AI-TT@simplelists.susmangodfrey.com>
Reply-To: Craig Smyser <Csmysyer@susmangodfrey.com>
Date: Friday, September 6, 2024 at 9:24 PM
To: "Wood, Emily Claire" <EWood@mofo.com>, "Ranieri, Vera" <VRanieri@mofo.com>, "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>, "MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright <OpenAICopyright@mofo.com>, "OpenAlcopyrightlitigation.lwteam@lw.com" <OpenAlcopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "sshoulder@cdas.com" <sshoulder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, Alter-TT <Alter-AI-TT@simplelists.susmangodfrey.com>
Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Emily,

We have set forth our position and understanding of the parties' meet and confers in my prior correspondence. Please provide a date next week by which OpenAI will provide the information, answers, and positions listed in my email of 8/27 and including a response to my correspondence of 7/29 regarding RFPs 34, 35, 45, 46, 47, 54, 56, 37, 50, 66, 55, and 61.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-2023 (office) | (713) 503 6252 (cell)
csmysyer@susmangodfrey.com | www.susmangodfrey.com

From: "Wood, Emily Claire" <EWood@mofo.com>

Date: Friday, August 30, 2024 at 4:31 PM

To: Craig Smyser <Csmysyer@susmangodfrey.com>, "Ranieri, Vera" <VRanieri@mofo.com>, "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>, "MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright <OpenAICopyright@mofo.com>, "OpenAICopyrightlitigation.lwteam@lw.com" <OpenAICopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com" <ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, Alter-TT <Alter-TT@simplelists.susmangodfrey.com>

Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Craig,

Given end of summer vacations, the upcoming holiday weekend, and the volume of RFPs at issue, we are not prepared to give Plaintiffs final positions on all the below requests by next Monday (which is a federal holiday) or Tuesday. We will get back to you with our positions in due course. In the meantime, please see below for OpenAI's summary of the parties' August 23rd meet and confer regarding OpenAI's responses to Plaintiffs' Third RFPs. Thank you and have a great holiday weekend.

RFP 58 [to the extent not already requested, documents relating to OpenAI's plans to allow creators to receive compensation or determine how and whether their works will be used by OpenAI, including all documents relating to Sam Altman's congressional testimony that "creators deserve control over how their creations are used"]

OpenAI expressed that a request for all documents relating to Sam Altman's congressional testimony is not relevant or proportional to the needs of the case. OpenAI noted that Plaintiffs have a copy of the testimony itself, and OpenAI stated that it is unclear how searching for all drafts or communications regarding the testimony would be proportional. In a follow up email sent on August 27, Plaintiffs requested OpenAI search for: (1) drafts or discussion of the phrase "creators deserve control over how their creations are used" in Sam Altman's congressional testimony and (2) your internal discussions regarding whether and how to compensate authors or rightsholders and/or enable opt-out of the use of

their data in training. OpenAI will consider this scope for the latter part of the request.

However, OpenAI noted that the former provision is incredibly broad, particularly in light of what OpenAI is already producing. OpenAI stated that it is not clear what Plaintiffs are searching for that would not be encompassed by other RFPs. Plaintiffs provided examples of documents Plaintiffs think are responsive to this RFP that would not be encompassed by prior requests, including documents discussing whether authors should be paid when their works are included in training data, and how—apart from robots.txt—creators should be able to control whether their works are used in OpenAI’s training data. OpenAI agreed to consider those specific requests, and Plaintiffs agreed that was a fine place to start.

RFP 59 [all documents relating to OpenAI’s decision(s) to obtain or request a license for training material from any copyright holder]

As an initial matter, OpenAI stated that it was not conceding that any of the documents at issue qualified as a license or a request for a license, but that that it would characterize these agreements as agreements relating to training data (although it was not withholding documents on the basis that OpenAI would characterize the documents that we understand Plaintiffs to be looking for differently than Plaintiffs do).

On the substance of the RFP, OpenAI stated that it does not understand how every document about an agreement is relevant or proportional to the case, and asked Plaintiffs to better explain what they are seeking with this request given OpenAI has already agreed to produce the agreements themselves. In Plaintiffs August 27 follow-up, Plaintiffs expressed that Plaintiffs believe the following categories of documents are responsive to Plaintiffs’ request: (1) discussions OpenAI had internally regarding its decision to enter into an agreement relating to training material from copyright holders, (2) drafts of the agreements, and (3) the agreements themselves (which Plaintiffs acknowledged OpenAI has already agreed to produce).

OpenAI reiterated that it was unclear why every document regarding OpenAI’s decision to enter a particular agreement is relevant or proportional to the case, particularly for agreements not related to commercial books. Plaintiffs stated that Plaintiffs position is that the market for training data is relevant to fair use and that Plaintiffs disagree that the data OpenAI receives related to works other than books is sufficiently different from Plaintiffs’ works to constitute a different market. Additionally, Plaintiffs stated that Plaintiffs have seen OpenAI productions thus far where OpenAI has approached the publisher with an in-kind rather than cash offer in exchange for the training data. Plaintiffs believe the decision to make an in-kind offer is relevant to the market for Plaintiffs’ works.

OpenAI agreed to consider Plaintiffs’ explanation.

RFP 60 [all documents describing the way in which OpenAI’s LLMs use training data, including human learning, pattern recognition, or parroting]

OpenAI stated that it was unclear what Plaintiffs mean by how LLMs “use training data.” Plaintiffs replied that it was Plaintiffs understanding that part of OpenAI’s argument for why the models are transformative is that the model learns from the underlying text like a human does—for example, by learning patterns, as OpenAI has allegedly stated publicly. So, Plaintiffs explained that Plaintiffs are searching for documents where OpenAI describes the models’ process in that manner because Plaintiffs believe that description bears on fair use. Plaintiffs also believe how the model is or is not creating novel

expression (as opposed to outputting information that summarizes or repeats the training data) is relevant.

OpenAI stated that Plaintiffs' second articulated rationale is likely covered by other RFPs regarding regurgitation. Plaintiffs disagreed because Plaintiffs believe other terms such as "parroting" would be responsive and do not refer to regurgitation outright. Plaintiffs stated that, to the extent OpenAI is representing that it would produce everything that relates to these terms in response to the RFPs it has served on regurgitation, Plaintiffs would consider that, but at the moment Plaintiffs are not sure all of these terms are covered.

OpenAI expressed that it was still not clear what Plaintiffs are searching for with the remainder of this request and asked for examples of responsive documents. Plaintiffs stated that Plaintiffs have seen public papers describing how LLMs learn, including some that use the term "parrot." Plaintiffs believe internal documents discussing those papers would be responsive, as would documents discussing whether LLMs learn the way humans learn, whether LLMs are "parrots," and whether LLMs learn through pattern recognition. OpenAI replied that how LLMs work is at the heart of this case and that OpenAI has likely produced many such relevant documents.

OpenAI stated that if this request is meant to target whether OpenAI internally uses the term "teaching" or "learning" to discuss how LLMs learn, it can consider searching for those documents. Plaintiffs asked OpenAI to take back that specific category of documents and stated that Plaintiffs will send a list of the public papers they referred to. If OpenAI agrees to search for documents relating to those two categories of documents, Plaintiffs may be able to accept that scope.

RFP 62 [to the extent not already requested, documents relating to the exhaustion of data available for training an LLM].

OpenAI asked Plaintiffs to define the meaning of "exhaustion" as used in this request. Plaintiffs explained that Plaintiffs have seen documents in which OpenAI executives have discussed whether there is a "data wall" and the possibility of running out of training data. OpenAI expressed that it was unclear why such documents are relevant to this case. Plaintiffs believe such documents are relevant to knowledge and willfulness, or to the relevant market for training data. OpenAI stated that it was not clear how these documents are relevant to knowledge and willfulness and that the connection to fair use seems tenuous and speculative, but that OpenAI understood what Plaintiffs were seeking with this request and would consider it.

RFP 63 [to the extent not already requested, all documents relating to OpenAI's decision to stop making public information about the training of OpenAI's LLMs]

OpenAI again expressed that it did not see how such documents were relevant. As an example, Plaintiffs explained that Plaintiffs believe OpenAI had debates about how to publicly describe the books data used in training. According to Plaintiffs, OpenAI ultimately decided to publicly label the books training data as books1/books2 and describe it very generically. Plaintiffs believe that decision and other choices about how much detail to publish about training data is relevant to knowledge and willfulness.

OpenAI stated that Plaintiffs' original RFP is broader than that description but that it would consider this narrowed scope if Plaintiffs are amenable. Plaintiffs asked OpenAI to consider rephrasing the request as one regarding "OpenAI's decision to stop making public information about the training **data** of OpenAI's

LLMs,” and opposed to only “training of OpenAI’s LLMs.” OpenAI stated that it would consider this request, though we still do not believe it is proportional, and would consider other ways to narrow the request.

RFP 65 [all documents relating to OpenAI’s use of web crawler permissions in connection with OpenAI’s efforts to comply with copyright law]

OpenAI expressed that documents relating to OpenAI’s efforts to comply with copyright law would be privileged and that it is not proportional to search for such documents. Plaintiffs expressed that Plaintiffs believe they have seen documents demonstrating that such discussions occurred across the company and not just amongst lawyers, so Plaintiffs believe responsive, non-privileged documents exist. And Plaintiffs disagree that the burden is disproportionate because Plaintiffs believe such documents are highly probative.

For example, Plaintiffs believe that efforts to design a web scraper that complies with OpenAI policies including those on copyright law would not be privileged. Plaintiffs expressed that they are not willing to narrow this request at this time and believe that, for example, documents from one engineer to another on this topic would be relevant and not privileged. OpenAI expressed that we do not believe that communication between engineers would even be responsive to this request because efforts to filter copyrighted information may or may not be an effort to comply with copyright law. Plaintiffs again noted that they disagree.

OpenAI expressed that the main issue with this request is the phrase “efforts to comply with copyright law.” For OpenAI to produce documents in response to this request, it would have to determine that a particular conversation is about complying with copyright law, which is privileged. And if Plaintiffs are merely seeking documents regarding OpenAI’s efforts to exclude copyrighted content, those documents are encompassed by other RFPs. Plaintiffs stated that Plaintiffs are not willing to narrow this request and will decide whether to move on it. OpenAI encouraged Plaintiffs to consider a compromise and to check existing RFPs to determine what non-privileged documents Plaintiffs believe are not encompassed by existing requests that would be responsive to this RFP. OpenAI also stated that we would consider ways to reframe this request that would not call for privileged documents.

RFP 67 [documents sufficient to show all prompts inputted into ChatGPT relating to any of the Class Works and the output ChatGPT generated in response to each prompt]

OpenAI expressed that we are not sure how we would begin searching for such documents. This is like asking Google to search through all its search queries—there is a huge volume of data and doing so raises PII concerns. And OpenAI is not sure how it would determine whether a prompt relates to a class work. OpenAI proposed tabling this discussion for now until OpenAI determines whether and how prompt and output data can be searched. Plaintiffs requested the parties consider this RFP in conjunction with related RFP 52 and asked to be permitted to inspect the prompt/output data if producing the data is unduly burdensome.

Plaintiffs asked OpenAI for an update on OpenAI’s efforts to determine whether prompt/output data is searchable by the week of August 26. OpenAI stated that it would pass along the request.

RFP 69 [all documents relating to whether OpenAI’s LLMs comprise a compression of training data]

OpenAI asked Plaintiffs whether there is a specific research theory Plaintiffs had in mind with this RFP. Plaintiffs stated that the terms “compression” and “compression of training data” are popular descriptions of training data and are used in various research papers. Plaintiffs offered to send the names of the referenced papers. OpenAI agreed to consider Plaintiffs’ request.

RFP 70 [all documents relating to the dataset known as The Pile and/or The Eye]

OpenAI asked Plaintiffs to explain why Plaintiffs need all documents on this subject, particularly when Plaintiffs can inspect OpenAI’s training data. Plaintiffs explained that Plaintiffs believe The Pile was designed to recreate the training dataset used for OpenAI’s models, and Plaintiffs believe all documents on this topic are relevant to the case. OpenAI stated that it was OpenAI’s understanding that The Pile is not a recreation of the *content* of an OpenAI dataset. If you have seen documents suggesting otherwise (whether publicly available or produced in this case), please provide them.

Plaintiffs believe that, even if OpenAI did not use The Pile for training, there could still be relevant discussions regarding the dataset. Plaintiffs believe the dataset is relevant because Plaintiffs believe OpenAI discussed whether or not to train on The Pile and whether The Pile is a superior dataset.

OpenAI asked Plaintiffs why such discussions are relevant, and Plaintiffs replied that they are relevant if OpenAI did download and use the dataset because doing so could be independent acts of copyright infringement. OpenAI expressed that it is still not clear why all documents about The Pile are relevant, but that OpenAI could consider a narrower request. Plaintiffs stated that if OpenAI is not willing to search for all documents relating to The Pile, the parties may be at an impasse on this issue. OpenAI disagreed that the parties were at an impasse and stated that OpenAI would get back to Plaintiffs with a final position. As written, OpenAI believes this request is too broad. OpenAI stated that it would consider ways to narrow the request and asked Plaintiffs to do the same.

From: Craig Smyser <CSmyser@susmangodfrey.com>

Sent: Tuesday, August 27, 2024 11:36 PM

To: Wood, Emily Claire <EWood@mofo.com>; Ranieri, Vera <VRanieri@mofo.com>; nyclassactions_microsoft_ohs@orrick.com; MicrosoftNYClassActionFDBR@faegredrinker.com; OpenAICopyright <OpenAICopyright@mofo.com>; OpenAICopyrightLitigation.lwteam@lw.com; KVPOAI@keker.com

Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley <wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; ssholder@cdas.com; ccole@cdas.com; Alter-AI-TT@simplelists.susmangodfrey.com

Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

External Email

Emily,

We disagree with numerous characterizations in your emails.

With respect to request 48, OpenAI is best suited to identify which Microsoft products incorporate its LLMs and/or ChatGPT. We expect OpenAI to produce for those products which it knows use OpenAI's LLMs and/or ChatGPT. We would at least expect it to include those products for which public reporting states that OpenAI's LLM's have been used, e.g., Bing (<https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/>), Edge (<https://www.microsoft.com/en-us/edge/features/ai?form=MA13FJ>), GitHub Copilot (<https://github.com/features/copilot>), Microsoft 365 Copilot (<https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/>), Power Platform (<https://techcommunity.microsoft.com/t5/apps-on-azure-blog/build-next-gen-apps-with-openai-and-microsoft-power-platform/ba-p/3742289>), Dynamics 365 CoPilot (<https://techcrunch.com/2023/03/06/microsoft-dynamics-copilot/>), and Azure OpenAI Service (https://learn.microsoft.com/en-us/azure/ai-services/openai/overview?WT.mc_id=academic-105485-koreyst). Please provide your response by next Tuesday.

With respect to request 49, 51, 52, and 53, please provide your response by next Tuesday.

With respect to request 57, we believe the request is fine as written given the large number of public statements by OpenAI leadership regarding these exact issues, e.g., <https://www.hbs.edu/bigs/sam-altman-harvard-business-school>, <https://openai.com/index/gpts-are-gpts/>, <https://arxiv.org/abs/2303.10130>, <https://fortune.com/2023/10/18/openai-sam-altman-misjudged-ai-creativity-chatgpt/>, <https://www.sfchronicle.com/bayarea/article/openai-funds-research-impact-universal-basic-19589117.php>, <https://www.techtimes.com/articles/304379/20240507/openai-ceo-sam-altman-admits-ai-eliminate-current-jobs.htm>, <https://openai.com/index/economic-impacts/>, <https://openai.com/global-affairs/testimony-of-sam-altman-before-the-us-senate/>, and <https://openai.com/index/language-model-safety-and-misuse/>. Please let us know if you will search for and produce documents responsive to this RFP by next Tuesday.

I further note that it does not appear OpenAI has responded to my correspondence of July 29th regarding RFPs 34, 35, 45, 46, 47, 54, 56, 37, 50, 66, 55, and 61. By next Tuesday, please provide confirmation that OpenAI will produce in response to these RFPs or that the parties are at an impasse so that we may consider seeking relief from the court. We note that it has been more than a month since OpenAI asked to meet and confer in its responses to Plaintiffs' Third RFPs and the parties have only just finished their initial discussion regarding Plaintiffs' Third RFPs.

With respect to RFP 68, on our meet and confer Vera said “[i]f there are

overlapping **issues**, we will produce the documents produced in the Tremblay case **even if you haven't served the exact RFPs.**" (emphasis added). I know Vera said this because I explicitly read the above sentence and asked if that was OpenAI's position on the matter, to which Vera agreed.

Your email suggests that OpenAI's position is now that OpenAI "will not produce documents responsive to a request by the Tremblay plaintiffs for specific information that has not been requested in this case, or documents specifically concerning the named plaintiffs in the Tremblay matter that are not otherwise relevant to this case."

First, there is no basis for OpenAI's shift in position from producing documents relating to overlapping "issues" to withholding documents if the "specific information [] has not been requested in this case." Please confirm that OpenAI's position is that it will produce the documents it is producing in *Tremblay* which relate to the overlapping issues in the two cases.

Second, given the overlap in issues and claims against OpenAI between the two cases, we do not believe there are any documents which OpenAI would produce in Tremblay "concerning the named plaintiffs in the Tremblay matter" that are "not otherwise relevant to [the SDNY case]." Please provide an example of a document which OpenAI would (or has) produced concerning an author plaintiff in Tremblay which would not be relevant to the instant case by next Monday.

We disagree with much of your summary, however, to simplify things, for RFPs 33, 38, 39, and 40, in which your email notes that OpenAI is considering the scope of the request, please provide a response by next Tuesday regarding OpenAI's position.

For RFPs 36, 42, 43, and 44, your email does not state that OpenAI is considering the scope or its production in response to these RFPs. Please confirm by next Tuesday that OpenAI is standing on its objection and that the parties are at an impasse.

For RFP 41, we believe that documents (1) demonstrating the techniques, technologies, or procedures, OpenAI employs to prevent its LLMs from displaying copyrighted material or regurgitating training data, and (2) showing the need for and development of the techniques, technologies, or procedures in (1), would at the very least be responsive to RFP 41. Depending on OpenAI's productions in response to, for non-exclusive examples, RFPs 24, 25, 27, and 28, these materials may already be produced. But, to the extent OpenAI has not or will not produce such documents in response to these or other RFPs, RFP 41 calls for their production.

The following concerns the meet and confer conducted on 8/23:

For RFP 58, you requested that we put categories in writing that you can pose to your client.

The categories are: (1) drafts or discussion of the phrase “creators deserve control over how their creations are used” in Sam Altman’s congressional testimony and (2) your internal discussions regarding whether and how to compensate authors or rightsholders and/or enable opt-out of the use of their data in training.

For RFP 59, we understood that you would ask your client regarding three categories specified on the call consisting of (1) discussions OpenAI had internally regarding its decision to obtain or seek a license for training material from copyright holders, (2) drafts of your license agreements, and (3) the agreements themselves (which I understand OpenAI is already producing). Please note that RFP 29 covers discussions with counterparties and potential counterparties. If you are not producing in response to either RFP, please let us know by next Tuesday.

For RFPs 60, 62, 63, and 69 I understood that you would be asking your client regarding these requests, please let us know what they say by next Tuesday.

I understood the parties agreed that RFP 67 should be paired with RFP 52 for resolution. I requested that you update us next week on OpenAI’s investigation of the feasibility of production or inspection of prompt-output data.

As to RFP 65 and 70, you objected to the scope of these requests and yet indicated that you would ask your client. If you have suggestions for a narrowed scope that would be acceptable to OpenAI, please let us know, otherwise we believe the requests as-written are appropriate.

Thanks,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-2023 (office) | (713) 503 6252 (cell)
csmysyer@susmangodfrey.com | www.susmangodfrey.com

From: "Wood, Emily Claire" <EWood@mofo.com>

Date: Monday, August 26, 2024 at 7:23 AM

To: Craig Smyser <CSmyser@susmangodfrey.com>, "Ranieri, Vera" <VRanieri@mofo.com>, "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>, "MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright <OpenAICopyright@mofo.com>, "OpenAICopyrightlitigation.lwteam@lw.com" <OpenAICopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier,

Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com" <ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, Alter-TT <Alter-TT@simplelists.susmangodfrey.com>

Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Craig,

Thanks for your email. Below is our summary of the meet and confer on August 16, regarding RFPs 48, 49, 51, 52, 53 and 57, as well as a response to your recent email on the same topic.

* * *

RFP 48 [all documents relating to the use or inclusion of OpenAI LLMs in Microsoft products, including Azure, and including documents relating to profits or revenues from such products and the distribution thereof among OpenAI and Microsoft]

OpenAI sought clarification regarding what information Plaintiffs are requesting, including what Plaintiffs meant about documents relating to the use or inclusion of ChatGPT in Azure. Plaintiffs explained that they understand Azure to be Microsoft's cloud computing infrastructure, and Plaintiffs believe it is used to service GPT or API endpoints. According to Plaintiffs, how Microsoft is arranging and charging for access to that infrastructure is relevant to the case. In response, OpenAI explained that it was difficult to identify a proportional set of documents to be produced if Plaintiffs' position is that Plaintiffs want all documents relating to the use of OpenAI LLMs in Microsoft products.

Plaintiffs then provided examples of categories of documents Plaintiffs believe are responsive. Plaintiffs stated that Plaintiffs would be satisfied if OpenAI produced those categories of documents in response to this RFP. Plaintiffs agreed to follow up to confirm the specific categories of documents Plaintiffs would accept in response to this request.

On August 20th, Plaintiffs proposed that they seek: (1) documents reflecting the decision to integrate ChatGPT or OAI's Large Language Models in Microsoft's Products, (2) documents reflecting the evaluation of Microsoft's Azure platform for serving / deploying ChatGPT or OAI's Large Language Models as part of Microsoft's Products and (3) documents showing the profits or revenues from the integration of ChatGPT or OAI's Large Language Models in Microsoft's Products. We consider this proposal to be Plaintiff's current position as to the documents it believes are necessary to consider the RFP satisfied. Can you provide the list of products that are encompassed by the term "Microsoft Products" or a definition of that term?

RFP 49 [all documents relating to the use or inclusion of OpenAI LLMs in commercial products, including documents relating to the ideation, development, and commercialization timeframe]

OpenAI stated that a request for all documents concerning OpenAI LLMs in commercial products is facially overbroad and that terms in this RFP, including "ideation," are vague. We accordingly requested Plaintiffs further describe what documents Plaintiffs are seeking with this request. Plaintiffs defined ideation to include "brainstorming, researching, and scoping." Plaintiffs explained that Plaintiffs are interested in documents describing any alleged changes OpenAI had to make to a product to render it appropriate for inclusion in some commercial product, along with documents about the development and ideation for that product. Specifically, Plaintiffs listed out broad categories of documents it seeks. OpenAI noted that it had concerns about burden and proportionality.

On August 20th, Plaintiffs proposed that OpenAI search for and produce: (1) documents showing the decision to integrate ChatGPT or OAI's Large Language Models in commercial products, (2) documents showing the development of the particularly commercial aspect of those products (as distinct from their general development as LLMs), and (3) documents showing the timeline / speed of development. We

consider this proposal to be Plaintiff's current position as to the documents it believes are necessary to consider the RFP satisfied.

OpenAI will consider Plaintiffs' request.

RFP 51 [the balance sheet, income statement, and cash flow statement, on a monthly basis, for each OpenAI entity]

OpenAI asked Plaintiffs to explain why—assuming they exist—Plaintiffs need each of these documents that are likely to provide cumulative information. Plaintiffs explained that Plaintiffs believe each of these documents are relevant to OpenAI's financial health. Plaintiffs also stated that they only want these documents in the manner in which they are kept in the ordinary course of business. For example, if OpenAI does not maintain monthly financial statements, Plaintiffs agreed it is not necessary for OpenAI to create monthly statements and can instead produce what is maintained in the ordinary course. Plaintiffs believe all information on OpenAI's balance sheets is relevant to OpenAI's financial health over time. Plaintiffs further explained that Plaintiffs believe OpenAI's financial health is relevant to disgorgement and provides context to OpenAI's decisions surrounding licensing, agreements and negotiations with Microsoft, and attempts to commercialize its products, but Plaintiffs did not explain why Plaintiffs need each likely duplicative document to understand those topics.

OpenAI will consider Plaintiffs' request.

RFP 52 [all documents sufficient to show the output of any OpenAI LLMs related to Plaintiffs' works]

OpenAI expressed that we were unsure how we would determine whether an output "relate[s] to Plaintiffs' works," and OpenAI noted that there could be an enormous amount of data that exists with respect to inputs and outputs generally. Plaintiffs asked if OpenAI would be willing to run a search for the name of Plaintiffs' works in the output data as a proposed resolution. OpenAI stated that it would consider that request, but that there could be contractual and privacy issues related to releasing output data as well. Plaintiffs also asked if Plaintiffs could inspect the output data if there is a burden issue with producing it. OpenAI stated that doing so could have its own burden issues if Plaintiffs are requesting OpenAI load a computer with every output in its possession and raises the same privacy concerns. Ultimately, OpenAI stated that it would take back the proposal of searching for the title of Plaintiffs' works in the output data, but also flagged that we were not sure such a search was even possible. On August 20th, Plaintiffs suggested that, "as an initial step," OpenAI search the output data for the named Plaintiffs' works and/or their names and, depending on the volume, either produce the data or make it available for inspection.

OpenAI will consider this request and investigate the possibility of running such searches and what is both technologically possible and reasonable.

RFP 53 [documents sufficient to show each version of the terms and/or conditions of use for ChatGPT]

OpenAI asked for a time-period for this request, and Plaintiffs replied that they are seeking these documents from the launch of ChatGPT (i.e., November 2022) to the present. OpenAI stated that it would consider this request.

RFP 57 [all documents concerning or relating to the impact or potential impact of OpenAI's LLMs on the markets for labor, art, or writing]

OpenAI stated that it was unsure how to define or create searches for terms such as a "market for labor." Plaintiffs replied that searching for the phrases "labor market," "art market," and "book market" is a start, and that Plaintiffs have seen public documents where OpenAI leadership has discussed the impact of AI on labor market transformation. Plaintiffs clarified that they are seeking documents about the broader impact of models on the livelihood of writers/workers and expressed that Plaintiffs believe such documents are relevant to factor 4 of the fair use analysis. Even assuming but not conceding Plaintiffs are correct as to relevance, OpenAI asked why it is proportional to seek all documents concerning or relating to these markets. Plaintiffs proposed a potential narrowed scope for this RFP where OpenAI would search for discussions about AI's labor impact or impact on the market for written works and any

research OpenAI has allegedly performed on those topics.

OpenAI asked if Plaintiffs would want to search for these documents across all custodians even though some individuals may not have expertise on this type of specialized topic. Plaintiffs expressed that regardless of the speaker's expertise, Plaintiffs believe a custodian's discussions about AI's impact on these markets could be relevant to knowledge and willfulness. Plaintiffs agreed to provide their proposal in writing, and OpenAI agreed to consider it. Plaintiffs did not provide a written proposal in their August 20th email. Please let OpenAI know if you intend to provide one. Please also provide the public documents you discussed during the meet and confer in connection with this RFP.

From: Wood, Emily Claire

Sent: Wednesday, August 21, 2024 10:28 AM

To: 'Craig Smyser' <CSmyser@susmangodfrey.com>; Ranieri, Vera <VRanieri@mofo.com>; nyclassactions_microsoft_ohs@orrick.com; MicrosoftNYClassActionFDBR@faegredrinker.com; OpenAICopyright <OpenAICopyright@mofo.com>; OpenAlcopyrightlitigation.lwteam@lw.com; KVPOAI@keker.com

Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley <wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; ssholder@cdas.com; ccole@cdas.com; Alter-AI-TT@simplelists.susmangodfrey.com

Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

Craig,

Please see below for OpenAI's summary of the parties' August 13 meet and confer regarding OpenAI's responses to Plaintiffs' Third RFPs, as well as a response to your email regarding RFP 68. OpenAI will respond to Plaintiffs' email regarding the parties' August 16 meet and confer shortly and is not agreeing with Plaintiffs' positions or summary by not addressing them in this email. Thank you.

* * *

Overarching Discussion

Plaintiffs expressed that they want OpenAI to provide Plaintiffs with a narrowed scope for any request where our response was an offer to meet and confer. We emphasized repeatedly that many of the requests Plaintiffs have served are broad or we do not understand what it is that they are seeking. Thus, it is Plaintiffs' burden to tell us what documents Plaintiffs want, particularly if they are seeking documents on a discrete topic or issue, so that we can assess overbreadth or how to locate these documents.

RFP 68 [all documents produced in Tremblay v. OpenAI, Inc., No. 23-cv-3223 (N.D. Cal.)]

OpenAI expressed our position that we will not withhold documents on the grounds that these documents were produced in Tremblay, if the documents are responsive to Plaintiff's RFPs and relevant in our case. So, we will produce a document that was also produced in *Tremblay* if it is responsive to Plaintiffs' RFPs, even if Plaintiffs have not served identical RFPs. We will not produce documents responsive to a request by the Tremblay plaintiffs for specific information that has not been requested in this case, or documents specifically concerning the named plaintiffs in the *Tremblay* matter that are not otherwise relevant to this case. Plaintiffs still need to serve proper requests for production. But we will

not withhold documents just because they were produced in *Tremblay* or another case if they are responsive here.

On the call, we asked if Plaintiffs were under the impression that we were withholding documents just because they were produced in *Tremblay*. Plaintiffs raised that OpenAI had previously inadvertently produced and then withdrawn interrogatory responses OpenAI meant to produce in the *Tremblay* matter. We stated that OpenAI's position is that we would object to a straight request for all discovery responses served in *Tremblay*.

Our understanding was that this resolved Plaintiffs' concerns, but Plaintiffs sent an email on Wednesday to further elaborate. In relevant part, the email says: "OpenAI may not withhold discovery on the basis that it relates to absent class members in the SDNY actions, regardless of their status as named plaintiffs in the *Tremblay* action." To explain your position, Plaintiffs provided an example of OpenAI's production of a complaint from a writer who is a named plaintiff in *Tremblay*, claiming that this should be produced in the instant case as well, even if that writer is not a named plaintiff.

The example Plaintiffs gave is not an example where OpenAI would withhold a document, assuming it would be responsive to some other RFP in the instant case to which OpenAI has agreed to produce documents. If a document is otherwise responsive to Plaintiffs' RFP and we've agreed to produce it in response to that RFP, it is not being withheld because it **also** relates to the named plaintiffs in the ND Cal. action. So, to the extent Plaintiffs are only seeking documents about "(for example) a class member who happens to be a named plaintiff in *Tremblay* **which are responsive either to request 68 or our other requests**," we have already agreed that they will not be withheld given those documents would be responsive to other requests (assuming we agreed to produce in response to one or more of those other requests). We do not believe there is actually a dispute here. However, we're willing to discuss if there's any remaining ambiguity.

RFP 33 [all documents regarding the deletion of any training datasets]

Plaintiffs expressed that documents regarding the deletion of any training datasets are relevant because, according to Plaintiffs, many of OpenAI's training datasets appear to have contained Plaintiffs' class works and, again according to Plaintiffs, to the extent OpenAI is not complying with its own data storage policies, the deletion is relevant. Plaintiffs also stated that identifying these datasets is relevant to the number of copies made of Plaintiffs' works and to willfulness. OpenAI noted that we already responded to an interrogatory on OpenAI's retention policies and that we are not sure what Plaintiffs are referring to by potential noncompliance with that policy, but that we would consider the scope of Plaintiffs' request.

RFP 36 [all communications with third parties related to the use or reproduction of copyrighted materials to train any AI model developed by OpenAI]

Plaintiffs stated that there are "lots of documents that would obviously be encompassed" by this request. According to plaintiffs, this would include documents about licensing, communications with Microsoft and other infrastructure providers regarding how OpenAI is storing copyrighted material, communications with third parties about using OpenAI technology, and guarantees or statements about reproduction. OpenAI noted that this request is expansive and illustrates our concern that Plaintiffs are serving catch-all RFPs that encompass such a broad swath of documents that we do not know what Plaintiffs are actually seeking, how it relates to claims, or how to search for the documents.

RFP 38 [all communications between you and any person from whom you obtained or sought to obtain a license for the use of text or audio in connection with training LLMs]

OpenAI expressed that we are already producing the agreements themselves and are not sure why the communications are relevant and why it is not overly burdensome to search for and produce them.

Plaintiffs stated that the communications are relevant to knowledge and willfulness, the fair use analysis (including the market for material), whether OpenAI had a valid license to some materials used in training, and whether the negotiations included terms limiting the use of the materials to particular

phases of training. On the final point, Plaintiffs explained that Plaintiffs believe it is relevant whether a particular source of content has more value at one phase of training vs. another. OpenAI said we would consider Plaintiffs' explanation of scope and relevance.

RFP 39 [all documents concerning or relating to curation and the type of data used to train LLMs, including documents concerning or relating to the importance of long stretches of text, of high-quality text, or of human-written text . . .]

OpenAI explained that we do not understand the terms, or know how to search for, documents related to "high-quality text" or "long stretches of text." Plaintiffs told OpenAI that the latter phrase appears in a paper written by OpenAI employees including Alec Radford and Ilya Sutskever titled "Improving Language Understanding by Generative Pre-Training." Based on Plaintiffs' review of public discussions regarding training data, Plaintiffs believe there are "obviously" internal documents regarding the importance or impact of the inclusion or exclusion of different types of data. Plaintiffs believe this information is relevant to the degree of transformation of the works and the fair use analysis, as well as the value OpenAI allegedly placed on particular types of data. OpenAI thanked Plaintiffs for the explanation and told Plaintiffs we would consider it.

RFP 40 [documents concerning or relating to the relative value of types of data for use in training LLMs]

Plaintiffs told OpenAI that the theory of relevance for this request is the same as for prior requests, including fair use factors 2 and 4, the value of data to the end model, and the general factual question of how models work and how they are trained on the data in question. Plaintiffs explained that Plaintiffs are looking for value both in qualitative and quantitative terms, that this RFP encompasses discussions about LLMs created by *others*, not just OpenAI, and that Plaintiffs are looking for discussions about all data, not just class works. OpenAI said we would consider this scope.

RFP 41 [to the extent not already requested, documents concerning or relating to modifying any parameters for tuning or restricting LLMs or ChatGPT from reproducing materials used in training]

Plaintiffs stated that these documents are relevant because they provide information on remedies and "to what extent these models in their base form will output copyrighted material." When we asked what Plaintiffs meant by "base models," Plaintiffs defined that term as models that have undergone pre-training, but not any post-training or fine-tuning to apply model specifications or policies to stop the model from infringing on copyright. OpenAI expressed that it is not clear what documents Plaintiffs are seeking with this RFP that would not be produced in response to other requests, and Plaintiffs agreed to go back and check Plaintiffs' list of RFPs to identify any gaps.

RFP 42 [documents concerning or relating to OpenAI creating human-generated text for use as training data for LLMs]

Plaintiffs framed this RFP as relevant to the importance and value of class works and copyrighted works of fiction and nonfiction more generally, as well as the question of OpenAI's ability to easily source substitutes for class works either by commissioning humans to create text for training or creating synthetic data. Plaintiffs clarified, however, that this RFP does not include synthetic data, but that documents that mention synthetic data could be responsive to it. For example, Plaintiffs speculated that, if OpenAI could easily generate synthetic data, OpenAI may not need to commission a large number of humans to create training data. Plaintiffs also clarified that the request is not about sourcing pre-existing data such as class works. Instead, Plaintiffs framed this request as specifically about commissioning humans to generate text to use as training data. As to the scope of this RFP, Plaintiffs also stated that this request encompasses discussions of human-generated content for non-OpenAI LLMs.

RFP 43 [all documents concerning or relating to changes in output quality of LLMs based on the inclusion or exclusion of human-generated text in the training data]

Plaintiffs stated that this request is relevant for the same reasons as the prior requests. OpenAI expressed that it was unclear to us what Plaintiffs mean by "output quality" and the definition of "human generated text" in this context. Plaintiffs explained that for output quality, Plaintiffs believe

OpenAI researches its LLMs' performance on a variety of metrics and that Plaintiffs are looking for documents regarding those assessments, including communications about and drafts of those assessments. But here, in contrast to the prior request, human generated content means any work created by a human, which would include pre-existing materials or class works. Plaintiffs explained that Plaintiffs need all these documents because, according to Plaintiffs, a question central to this case is what value the class works or similar works have to the construction, improvement, or training of LLMs, and Plaintiffs believe that OpenAI has performed research on the optimal mix of training data and how to source sufficient training data.

RFP 44 [all documents concerning or relating to ChatGPT or LLM commercial uses and/or applications]

Plaintiffs explained this request as relevant to the fair use analysis and the market for the works in question. OpenAI explained that we were struggling to identify a limiting principle here as to what OpenAI documents are not encompassed by this request. We asked if the term "commercial" modified both "uses" and "applications," and Plaintiffs confirmed that the RFP should be read as requesting documents concerning commercial uses and/or commercial applications. OpenAI then asked why Plaintiffs need all documents on these subjects, and Plaintiffs explained that the RFP is "broad because of the nature of the case." Plaintiffs emphasized that there are questions surrounding whether, according to Plaintiffs, the class works help OpenAI profit off its LLMs and how OpenAI's models effect the market for the class works.

From: Craig Smyser <CSmyser@susmangodfrey.com>

Sent: Tuesday, August 20, 2024 10:42 AM

To: Ranieri, Vera <VRanieri@mofo.com>; Wood, Emily Claire <EWood@mofo.com>; nyclassactions_microsoft_ohs@orrick.com; MicrosoftNYClassActionFDBR@faegredrinker.com; OpenAICopyright <OpenAICopyright@mofo.com>; OpenAIcopyrightlitigation.lwteam@lw.com; KVPOAI@keker.com

Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley <wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; ssholder@cdas.com; ccole@cdas.com; Alter-AI-TT@simplelists.susmangodfrey.com

Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

External Email

Vera,

Thanks for meeting on Friday, please provide times this week to confer over requests 58, 59, 60, 62, 63, 65, 67, 69, and 70.

With respect to request 48, we propose that OpenAI search for and produce (1) documents reflecting the decision to integrate ChatGPT or OAI's Large Language Models in Microsoft's Products, (2) documents reflecting the evaluation of Microsoft's Azure platform for serving / deploying ChatGPT or OAI's Large Language Models as part of Microsoft's Products and (3) documents showing the profits or revenues from the integration of ChatGPT or OAI's Large Language Models in Microsoft's Products. Please let us know your position on these three

categories.

With respect to request 49, we propose that OpenAI search for and produce (1) documents showing the decision to integrate ChatGPT or OAI's Large Language Models in commercial products, (2) documents showing the development of the particularly commercial aspect of those products (as distinct from their general development as LLMs), and (3) documents showing the timeline / speed of development. Please let us know your position on these three categories.

With respect to request 52, we suggest that, as an initial step, you search the output data for the named plaintiff's works and/or their names. Depending on the volume which results, OpenAI could either produce this data or make it available for inspection as part of the training data / model inspection. Please let us know your position on this.

With respect to request 51, you were going to ask your client. With respect to request 53, we specified that we wanted the terms and conditions from the launch of ChatGPT to the present.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-2023 (office) | (713) 503 6252 (cell)
csmysersusmangodfrey.com | www.susmangodfrey.com

From: <Alter-AI-TT@simplelists.susmangodfrey.com> on behalf of Craig Smyser
<Csmysersusmangodfrey.com>

Reply-To: Craig Smyser <Csmysersusmangodfrey.com>

Date: Wednesday, August 14, 2024 at 12:11 PM

To: "Ranieri, Vera" <VRanieri@mofo.com>, "Wood, Emily Claire" <EWood@mofo.com>, "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>, "MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright

<OpenAICopyright@mofo.com>, "OpenAICopyrightlitigation.lwteam@lw.com" <OpenAICopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>

Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com" <ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, "Alter-AI-TT@simplelists.susmangodfrey.com" <Alter-AI-TT@simplelists.susmangodfrey.com>

Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Vera,

Thanks, we have sent the invitation.

Regarding yesterday's discussion on the meet and confer regarding Plaintiffs' RFP No. 68, we understand OpenAI's position to be that, for overlapping issues with the *Tremblay* matter, OpenAI will produce the documents produced in the *Tremblay* case even if the SDNY Plaintiffs have not served the exact RFPs. We appreciate that clarification.

You had also asked our position regarding documents which OpenAI may produce regarding the named plaintiffs in *Tremblay*. We understood OpenAI's position to be that while it would produce documents relating to class members (including absent class members) generally, it would not produce documents regarding the named plaintiffs in *Tremblay* on the grounds that the SDNY plaintiffs were not producing discovery regarding absent class members.

To the extent this represents OpenAI's position, it is improper. OpenAI may not withhold discovery on the basis it relates to absent class members in the SDNY actions, regardless of their status as named plaintiffs in the *Tremblay* action. To provide a concrete example, if, in *Tremblay*, OpenAI is producing a complaint from a writer who happens to be a named plaintiff in *Tremblay* about unlawful ingestion of their copyright protected book, that should be produced in the SDNY actions even if the writer is not a named plaintiff in the SDNY actions.

This situation is entirely different from whether OpenAI may propound discovery on absent class members; OpenAI may not absent the sort of good cause that does not exist here.

Please confirm that, if OAI has documents about (for example) a class member who happens to be a named plaintiff in *Tremblay* which are responsive either to request 68 or our other requests, they will not be withheld.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-2023 (office) | (713) 503 6252 (cell)
csmysersusmangodfrey.com | www.susmangodfrey.com

From: "Ranieri, Vera" <VRanieri@mofo.com>

Date: Tuesday, August 13, 2024 at 3:46 PM

To: Craig Smyser <Csmysersusmangodfrey.com>, "Wood, Emily Claire" <EWood@mofo.com>, "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>, "MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright

<OpenAICopyright@mofo.com>, "OpenAICopyrightlitigation.lwteam@lw.com"
<openaicopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com" <ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, "Alter-AI-TT@simplelists.susmangodfrey.com" <Alter-AI-TT@simplelists.susmangodfrey.com>
Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Hi Craig,

11 am PT on Friday would work for us, if you'd like to circulate an invite.

Thanks,

Vera

From: Craig Smyser <CSmyser@susmangodfrey.com>
Sent: Tuesday, August 13, 2024 12:10 PM
To: Wood, Emily Claire <EWood@mofo.com>; nyclassactions_microsoft_ohs@orrick.com; MicrosoftNYClassActionFDBR@faegredrinker.com; OpenAICopyright <OpenAICopyright@mofo.com>; OpenAICopyrightlitigation.lwteam@lw.com; KVPOAI@keker.com
Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley <wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; ssholder@cdas.com; ccole@cdas.com; Alter-AI-TT@simplelists.susmangodfrey.com
Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

External Email

Counsel:

Please let us know your availability to continue this meet and confer to discuss requests 49, 51, 52, 53, 57, 58, 59, 60, 62, 63, 65, 67, 69, and 70 later this week.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-2023 (office) | (713) 503 6252 (cell)
csmysyer@susmangodfrey.com | www.susmangodfrey.com

From: Craig Smyser <CSmyser@susmangodfrey.com>
Date: Monday, August 12, 2024 at 11:36 AM
To: "Wood, Emily Claire" <EWood@mofo.com>, "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>, "MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright <OpenAICopyright@mofo.com>, "OpenAICopyrightlitigation.lwteam@lw.com" <OpenAICopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com" <ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, "Alter-AI-TT@simplelists.susmangodfrey.com" <Alter-AI-TT@simplelists.susmangodfrey.com>
Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

Emily,

OpenAI is the party resisting discovery and which initially offered to meet and confer on requests 33, 36, 38, 39, 40, 41, 42, 43, 44, 48, 49, 51, 52, 53, 57, 58, 59, 60, 62, 63, 65, 67, 69, and 70. Given that—and that it has been nearly a month since we requested a meet and confer on these requests--OpenAI's refusal to provide its position on the appropriate scope of these requests is inappropriate. We request that you come prepared with your positions to the call on Tuesday.

We will send the invitation.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.
One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-2023 (office) | (713) 503 6252 (cell)
csmyser@susmangodfrey.com | www.susmangodfrey.com

From: "Wood, Emily Claire" <EWood@mofo.com>
Date: Sunday, August 11, 2024 at 9:31 PM
To: Craig Smyser <CSmyser@susmangodfrey.com>, "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>, "MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright <OpenAICopyright@mofo.com>, "OpenAICopyrightlitigation.lwteam@lw.com"

<openaicopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com" <ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, "Alter-AI-TT@simplelists.susmangodfrey.com" <Alter-AI-TT@simplelists.susmangodfrey.com>
Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Craig,

We acknowledge your request for OpenAI's position on the RFPs you mention below, but we need to better understand what the Times is looking for before we can provide our position on the proper scope of the requests. We are happy to discuss on our upcoming call.

Please send the invitation to the Latham, Keker, and MoFo listservs on this chain. Thank you.

Best,
Emily

From: Craig Smyser <CSmyser@susmangodfrey.com>
Sent: Friday, August 9, 2024 12:49 PM
To: Wood, Emily Claire <EWood@mofo.com>; nyclassactions_microsoft_ohs@orrick.com; MicrosoftNYClassActionFDBR@faegredrinker.com; OpenAICopyright <OpenAICopyright@mofo.com>; OpenAICopyrightlitigation.lwteam@lw.com; KVPOAI@keker.com
Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley <wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; ssholder@cdas.com; ccole@cdas.com; Alter-AI-TT@simplelists.susmangodfrey.com
Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

External Email

Emily,

Thanks, we can be available on Tuesday from 2-3 pm ET. Please provide OpenAI's position on the requests prior to the meeting and please let me know who should receive the invitation.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-2023 (office) | (713) 503 6252 (cell)
csmysyer@susmangodfrey.com | www.susmangodfrey.com

From: "Wood, Emily Claire" <EWood@mofo.com>
Date: Thursday, August 8, 2024 at 11:39 PM
To: Craig Smyser <CSmyser@susmangodfrey.com>,
<nyclassactions_microsoft_ohs@orrick.com> <nyclassactions_microsoft_ohs@orrick.com>,
<MicrosoftNYClassActionFDBR@faegredrinker.com>
<MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright
<OpenAICopyright@mofo.com>, <OpenAlcopyrightlitigation.lwteam@lw.com>
<openaicopyrightlitigation.lwteam@lw.com>, <KVPOAI@keker.com> <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier,
Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, <ssholder@cdas.com>
<ssholder@cdas.com>, <ccole@cdas.com> <ccole@cdas.com>, <[Alter-AI-TT@simplelists.susmangodfrey.com](mailto:Alter-AI-
TT@simplelists.susmangodfrey.com)> <Alter-AI-TT@simplelists.susmangodfrey.com>
Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI -
SDNY No. 23-10211

EXTERNAL Email

Hi Craig,

OpenAI is available to continue this meet and confer next Tuesday, August 13 from 2-3pm ET or
Thursday, August 15 from 3-4pm ET. Thank you.

Best,
Emily

From: Craig Smyser <CSmyser@susmangodfrey.com>
Sent: Thursday, August 8, 2024 10:42 AM
To: Wood, Emily Claire <EWood@mofo.com>; nyclassactions_microsoft_ohs@orrick.com;
MicrosoftNYClassActionFDBR@faegredrinker.com; OpenAICopyright <OpenAICopyright@mofo.com>;
OpenAlcopyrightlitigation.lwteam@lw.com; KVPOAI@keker.com
Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley
<wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; <ssholder@cdas.com>; ccole@cdas.com; [Alter-AI-TT@simplelists.susmangodfrey.com](mailto:Alter-
AI-TT@simplelists.susmangodfrey.com)
Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No.
23-10211

External Email

Following up on the below. It has been a week since you promised to follow-up with

proposed times. Please provide times for a meet and confer and OpenAI's position on the outstanding requests listed in my last email.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-2023 (office) | (713) 503 6252 (cell)
csmysyer@susmangodfrey.com | www.susmangodfrey.com

From: Craig Smyser <csmysyer@susmangodfrey.com>
Date: Friday, August 2, 2024 at 11:15 AM
To: "Wood, Emily Claire" <EWood@mofo.com>, "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>, "MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright <OpenAICopyright@mofo.com>, "OpenAICopyrightlitigation.lwteam@lw.com" <OpenAICopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com" <ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, "Alter-AI-TT@simplelists.susmangodfrey.com" <Alter-AI-TT@simplelists.susmangodfrey.com>
Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

Emily,

We are amenable to that approach but given the amount of time that has passed since our first request for a meet and confer please also provide OpenAI's position on the appropriate scope for request 68 and requests 33, 36, 38, 39, 40, 41, 42, 43, 44, 48, 49, 51, 52, 53, 57, 58, 59, 60, 62, 63, 65, 67, 69, and 70 in the email early next week with proposed times.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-2023 (office) | (713) 503 6252 (cell)
csmysyer@susmangodfrey.com | www.susmangodfrey.com

From: "Wood, Emily Claire" <EWood@mofo.com>
Date: Thursday, August 1, 2024 at 5:42 PM
To: Craig Smyser <CSmyser@susmangodfrey.com>,
"nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>,
"MicrosoftNYClassActionFDBR@faegredrinker.com"
<MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright
<OpenAICopyright@mofo.com>, "OpenAICopyrightlitigation.lwteam@lw.com"
<openaicopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com"
<ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, "Alter-AI-TT@simplelists.susmangodfrey.com" <Alter-AI-TT@simplelists.susmangodfrey.com>
Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Craig,

Thank you for your email. We are considering the caselaw Plaintiffs provided and, in the interest of addressing these discovery issues all at once, we propose continuing our meet and confer when we have had the chance to review the relevant caselaw and confer with our client. We will follow up early next with proposed times. Thank you.

Best,
Emily

From: Craig Smyser <CSmyser@susmangodfrey.com>
Sent: Monday, July 29, 2024 12:17 PM
To: Wood, Emily Claire <EWood@mofo.com>; nyclassactions_microsoft_ohs@orrick.com;
MicrosoftNYClassActionFDBR@faegredrinker.com; OpenAICopyright <OpenAICopyright@mofo.com>;
OpenAICopyrightlitigation.lwteam@lw.com; KVPOAI@keker.com
Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley <wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; ssholder@cdas.com; ccole@cdas.com; Alter-AI-TT@simplelists.susmangodfrey.com
Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

External Email

Thanks for speaking last Friday.

Please let us know when you can be available for meet and confer on the additional requests which we did not discuss early this week. By my count that would be request 68,

and then requests 33, 36, 38, 39, 40, 41, 42, 43, 44, 48, 49, 51, 52, 53, 57, 58, 59, 60, 62, 63, 65, 67, 69, and 70. It would be helpful if OpenAI would provide its position on the appropriate scope for these requests prior to the meet and confer.

Regarding Requests 34 and 35, you had asked us to send case law on the discoverability of retention agreements in light of the deletion of Books1 and Books2. *See, e.g., Linet Americas, Inc. v. Hill-Rom Holdings, Inc.*, 2023 WL 9119836, at *1-*3 (N.D.Ill., 2023) (deletion of relevant documents was a sufficient factual basis to permit discovery of the "retention letter and a log containing the dates, participants in, and subject matter of communications between Linet and its counsel related to potential claims against Hillrom from December 2019 to April 2020"). The scope of that retention is not privileged. *Id.*; *Berall v. Teleflex Medical Incorporated*, 2021 WL 5989936, at *2 (S.D.N.Y., 2021).

Regarding request 45 (documents relating to conversion to for-profit)*, you were going to confirm with your client that their position remains the same as in the *New York Times* case. We will review the briefing from the *New York Times* case on this issue and letting you know our position on the parties mutually agreeing to be bound by the Court's eventual decision on that motion.

Regarding request 46 (documents regarding whether your use of copyrighted material in LLMs complies with copyright law in the United States or elsewhere), you will look at the current search terms to determine your view on whether the relevant non-privileged documents responsive to this request would already be captured and, to the extent they were not, propose an alternative scope as to the portion of the request dealing with compliance with United States Copyright law. You had asked if we had case law supporting the relevance of whether OpenAI's use of copyrighted material complies with non-US Copyright laws. Please *see Bluewater Music Servs. Corp. v. Spotify USA Inc.*, No. 3:17-CV-01051, 2018 WL 10152456, at *1, *8-9 (M.D. Tenn. July 10, 2018) (ordering production of documents showing that Defendant was "able to appropriately pay royalties for the same musical compositions to music rights organizations in other countries" over Defendant's objection that compositions by entities "operating under foreign laws, would [not] put Defendant on notice of its alleged non-compliance with U.S. copyright law . . .").

Regarding request 47 (documents regarding products, services, or "in-kind" investments provided to OpenAI by Microsoft and the importance of the same), you had asked us to propose a narrowed scope. Plaintiffs are not interested in whether OpenAI is using Microsoft Word as its default word processing software or OpenAI's general usage of Microsoft's products. To the extent OpenAI had interpreted the request as including such documents, we agree to narrow the scope. Plaintiffs rather seek the documents regarding the products, services, or "in-kind" investments provided by Microsoft specifically for OpenAI's development, deployment, commercialization, and training of large language models, as well as the importance of the same. Please let us know if OpenAI will produce these documents.

Regarding request 54 (documents concerning the addition of certain language to OpenAI's terms and conditions for ChatGPT), you had asked us to identify the relevance of this material and consider a revised scope. As to a revised scope, we believe the request is appropriate as written.

As to relevance, OpenAI's terms and conditions state that "By uploading any User Content you hereby grant and will grant OpenAI and its affiliated companies a nonexclusive, worldwide, royalty free, fully paid up, transferable, sublicensable, perpetual, irrevocable license to copy, display, upload, perform, distribute, store, modify and otherwise use your User Content **for any OpenAI-related purpose** in any form, medium or technology now known or later developed." As stated on the call, we believe that OpenAI could claim to be authorized to, e.g., train its models on such user submitted content. We further note that OpenAI claims exactly that on its website:

"Will OpenAI use files uploaded to train its models?"

The answer depends on the service you are using. As explained in [this article](#), we may use content submitted to ChatGPT, DALL·E, and our other services for individuals to improve model performance. Content may include files that are uploaded. Please refer to this [article](#) to understand how content may be used to improve model performance and the choices that users have."

https://help.openai.com/en/articles/8555545-file-uploads-faq#h_3137baa34c

That is relevant to this case as it may be an additional source of Plaintiffs' copyrighted material used in training OpenAI's models.

Regarding request 56 (documents regarding "Fair Learning" by Mark Lemley), I understood that you were going to ask your team/client regarding whether OpenAI would search for and produce in response to this request and let us know.

Regarding request 37 (documents regarding the preparation of OpenAI's comment to the USPTO), your position was that OpenAI would be willing to search for and produce in response to this request if Plaintiffs' did the same regarding certain RFPs directed to their statements regarding AI. Please let me know if that accurately portrays OpenAI's position, and, if so, which specific RFPs to Plaintiffs are being referenced.

Regarding request 50 (financial information by month), I understood that you were going to speak with your client regarding the "cadence" at which the requested financial records are kept and let us know regarding the possession of these records regarding a certain subset of OpenAI's LLM products/services. I understood that your response to us on this issue would identify the LLM products/services in question.

Regarding request 66 (documents relating to media manager and its features), I understood

your position to be that OpenAI would search for non-privileged documents on this issue.

Regarding request 55 (documents regarding repositories of text used or considered using in training OpenAI's LLMs), as to versions of repositories of text used in training OpenAI's Large Language Models, this information is clearly relevant to Plaintiffs' claims as, for instance, determining which edition of LibGen was downloaded is necessary to determine which of Plaintiffs' works are contained within and thus used in training. The same is true of other repositories as, for example, the particular copyrighted works included in CommonCrawl, posted on Reddit, downloaded from Open Library, etc. are directly relevant to Plaintiffs' claims. The same is true for repositories of text considered but not used as, for example, discussions regarding which repositories are worth (or not worth) using in training, is directly relevant to the value of the repositories actually used in training the models. For example, in considering how to enable certain capabilities in its models, OpenAI's employees appear to have had discussions regarding the value of—and whether to use—various sources of text data.

Regarding request 61 (documents showing all audiobooks OpenAI transcribed), you had asked us to further explain the relevance for this request. As to audiobooks included in training its Large Language Models (and putting aside any dispute as to which models that phrase covers), this information is clearly relevant to Plaintiffs' claims and to our ability to effectively inspect the training data OpenAI is making available. As to other audiobooks, we note that such information is relevant at least to OpenAI's degree of knowledge and willfulness and, to the extent such books were used to, e.g. optimize other datasets used in training or as part of datasets used in evaluation rather than training, they are relevant to Plaintiffs' claims. They may also be evidence of direct infringement.

Thanks,
Craig

*Parentheticals have been provided for ease of reference in correspondence, requests as-written and served otherwise control.

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
 (212) 336-8330 (office) | (713) 503 6252 (cell)
csmysyer@susmangodfrey.com | www.susmangodfrey.com

From: "Craig Smyser (via Alter-AI-TT list)" <Alter-AI-TT@simplelists.susmangodfrey.com>

Reply-To: Craig Smyser <CSmyser@susmangodfrey.com>

Date: Tuesday, July 23, 2024 at 6:01 PM

To: "Wood, Emily Claire" <EWood@mofo.com>, "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>,

"MicrosoftNYClassActionFDBR@faegredrinker.com"
<MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright
<OpenAICopyright@mofo.com>, "OpenAlcopyrightlitigation.lwteam@lw.com"
<openaicopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com"
<ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, "Alter-AI-TT@simplelists.susmangodfrey.com" <Alter-AI-TT@simplelists.susmangodfrey.com>
Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Emily,

Thanks, we can be available this Friday from 9-10 AM PT. Please include myself, Rachel Geman, Rohit Nath, and Alejandra Salinas on the invitation.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-8330 (office) | (713) 503 6252 (cell)
csmysyer@susmangodfrey.com | www.susmangodfrey.com

From: "Wood, Emily Claire" <EWood@mofo.com>

Date: Tuesday, July 23, 2024 at 5:01 PM

To: Craig Smyser <CSmyser@susmangodfrey.com>,
"nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>,
"MicrosoftNYClassActionFDBR@faegredrinker.com"
<MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright
<OpenAICopyright@mofo.com>, "OpenAlcopyrightlitigation.lwteam@lw.com"
<openaicopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com"
<ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, "Alter-AI-TT@simplelists.susmangodfrey.com" <Alter-AI-TT@simplelists.susmangodfrey.com>

Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Hi Craig,

OpenAI is available to meet and confer this Friday, July 26 between 9-10am or 12-2pm PT. Please let me

know what time works best and I will send an invite. Thank you.

Best,
Emily

From: Craig Smyser <CSmyser@susmangodfrey.com>
Sent: Tuesday, July 23, 2024 10:11 AM
To: Wood, Emily Claire <EWood@mofo.com>; nyclassactions_microsoft_ohs@orrick.com; MicrosoftNYClassActionFDBR@faegredrinker.com; OpenAICopyright <OpenAICopyright@mofo.com>; OpenAlcopyrightlitigation.lwteam@lw.com; KVPOAI@keker.com
Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley <wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; ssholder@cdas.com; ccole@cdas.com; Alter-AI-TT@simplelists.susmangodfrey.com
Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

External Email

Emily,

Following up on the below, please let us know times this week for a meet and confer on these issues.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-8330 (office) | (713) 503 6252 (cell)
csmyser@susmangodfrey.com | www.susmangodfrey.com

From: "Wood, Emily Claire" <EWood@mofo.com>
Date: Wednesday, July 17, 2024 at 6:08 PM
To: Craig Smyser <CSmyser@susmangodfrey.com>,
"nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>,
"MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright
<OpenAICopyright@mofo.com>, "OpenAlcopyrightlitigation.lwteam@lw.com" <OpenAlcopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>
Cc: "Dupree, Andrea" <ADupree@mofo.com>, "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "ssholder@cdas.com"

<ssholder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, "Alter-AI-TT@simplelists.susmangodfrey.com" <Alter-AI-TT@simplelists.susmangodfrey.com>

Subject: RE: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Craig,

We are reviewing your email and will get back to you in due course. We are not available to meet confer on these issues this week but will get back to you next week with proposed times. Thanks.

Best,

Emily

From: Craig Smyser <CSmyser@susmangodfrey.com>

Sent: Tuesday, July 16, 2024 10:26 PM

To: nyclassactions_microsoft_ohs@orrick.com; MicrosoftNYClassActionFDBR@faegredrinker.com;

OpenAICopyright <OpenAICopyright@mofo.com>; OpenAlcopyrightlitigation.lwteam@lw.com;

KVPOAI@keker.com

Cc: Dupree, Andrea <ADupree@mofo.com>; Geman, Rachel <rgeman@lchb.com>; Dozier, Wesley <wdozier@lchb.com>; Stoler, Reilly T. <rstoler@lchb.com>; ssholder@cdas.com; ccole@cdas.com; Alter-AI-TT@simplelists.susmangodfrey.com

Subject: Re: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

External Email

Counsel,

Regarding your invitation to meet and confer concerning requests 33, 36, 38, 39, 40, 41, 42, 43, 44, 48, 49, 51, 52, 53, 57, 58, 59, 60, 62, 63, 65, 67, 69, and 70, we can be available for a meet and confer on these requests and the below issues this Thursday from 11 AM PT to 2 PM PT. In the interest of efficiency, it would be a great help if you would submit any proposals you have regarding any narrowed scope for these requests in writing in advance of that discussion. At that time please also be prepared to discuss the below issues:

For requests 34, 35, 45, 46, 47, 54, and 56, to which you have refused to produce any documents, we disagree with your objections and positions, at least for the following reasons:

- Requests 34 and 35, directly relate to the issue of when OpenAI was aware of potential copyright litigation and thus should have been affirmatively taking steps to

preserve documents.

- For request 45, OpenAI’s contemplation of for-profit status bears on OpenAI’s fair use defense.
- For request 46, OpenAI itself has taken the position that compliance with copyright laws is a must-have for its models and that it will indemnify end users against accusations of copyright infringement; there is no basis for OpenAI to refuse to produce documents concerning its alleged compliance with United States Copyright Law. *See, e.g.*, <https://cdn.openai.com/spec/model-spec-2024-05-08.html#respect-creators-and-their-rights>; <https://openai.com/policies/service-terms/> (Section 3.b).
- For request 47, Microsoft’s involvement in the training and development of the models at issue in this suit, including through in-kind investments of compute and expertise (touted repeatedly by leadership of both companies, *see, e.g.*, <https://www.nytimes.com/2024/07/14/technology/microsoft-ai-satya-nadella.html>; <https://openai.com/index/openai-and-microsoft-extend-partnership/> (“We’ve worked together to build multiple supercomputing systems powered by Azure, which we use to train all of our models. Azure’s unique architecture design has been crucial in delivering best-in-class performance and scale for our AI training and inference workloads.”)).
- For request 54, documents related to the inclusion, modification, or removal, of the copyright provision for the relevant terms and conditions are obviously relevant, at least to willfulness and to OpenAI’s understanding of its copyright obligations.
 - The same is true as to request 56.

Regarding your response to request 37, the comment itself is insufficient, OpenAI should produce the drafts, discussions, and other input documents which preceded the comment.

Regarding your response to request 50, your response states that you will produce documents sufficient to show “income from paid versions of ChatGPT.” Please confirm that you will produce documents sufficient to show the same for all of the Large Language Models which OpenAI has commercialized, sold, or licensed. Additionally, please confirm that you will produce monthly gross revenues, monthly net revenues, monthly profits, and licensing fees in addition to any other kinds of revenues and profits from ChatGPT and OpenAI’s commercialized/sold/licensed Large Language Models.

Regarding your response to request 66, again the description of the function of the particular tool at the link you cited is insufficient, the documents surrounding the tool’s creation, design, deployment, and implementation are also relevant.

Regarding your response to request 55 and 61, inspection of training data is insufficient. The sources OpenAI considered or evaluated training the publicly available models on, including pirated book repositories or its library of transcribed audiobooks, and the discussions which OpenAI conducted concerning such sources are relevant and should be

produced.

Regarding your response to request 68, your position is untenable given your position on, e.g., discovery and deposition coordination between the *Tremblay* action and the instant action.

Best,
Craig

J. Craig Smyser | Susman Godfrey L.L.P.

One Manhattan West, 51st Floor, New York, NY 10001
(212) 336-8330 (office) | (713) 503 6252 (cell)
csmysyer@susmangodfrey.com | www.susmangodfrey.com

From: "Dupree, Andrea (via Alter-AI-TT list)" <Alter-AI-TT@simplelists.susmangodfrey.com>

Reply-To: "Dupree, Andrea" <ADupree@mofo.com>

Date: Monday, July 15, 2024 at 6:07 PM

To: "Geman, Rachel" <rgeman@lchb.com>, "Dozier, Wesley" <wdozier@lchb.com>, "Stoler, Reilly T." <rstoler@lchb.com>, "sshoulder@cdas.com" <sshoulder@cdas.com>, "ccole@cdas.com" <ccole@cdas.com>, "Alter-AI-TT@simplelists.susmangodfrey.com" <Alter-AI-TT@simplelists.susmangodfrey.com>

Cc: "nyclassactions_microsoft_ohs@orrick.com" <nyclassactions_microsoft_ohs@orrick.com>, "MicrosoftNYClassActionFDBR@faegredrinker.com" <MicrosoftNYClassActionFDBR@faegredrinker.com>, OpenAICopyright <OpenAICopyright@mofo.com>, "OpenAlcopyrightlitigation.lwteam@lw.com" <OpenAlcopyrightlitigation.lwteam@lw.com>, "KVPOAI@keker.com" <KVPOAI@keker.com>

Subject: SERVICE: Authors Guild v. OpenAI, Inc. - SDNY No. 23-08292 & Alter v. OpenAI - SDNY No. 23-10211

EXTERNAL Email

Dear counsel,

Attached for service please find the following document:

- DEFENDANTS' RESPONSES AND OBJECTIONS TO PLAINTIFFS' THIRD SET OF REQUESTS FOR PRODUCTION OF DOCUMENTS

Thank you.

Sincerely,

ANDRÉA DUPRÉE

Legal Secretary

ADupree@mofo.com

T +1 (415) 268-6239



mfo.com | [LinkedIn](#) | [Twitter](#)

=====

====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

To unsubscribe from this list please go to <https://simplelists.susmangodfrey.com/subs/>

=====

====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

To unsubscribe from this list please go to <https://simplelists.susmangodfrey.com/subs/>

=====

====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

=====

====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

=====

====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

=====

====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

To unsubscribe from this list please go to <https://simplelists.susmangodfrey.com/subs/>

=====

====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

=====

====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

To unsubscribe from this list please go to <https://simplelists.susmangodfrey.com/subs/>

=====

====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

=====

====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

=====

====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

To unsubscribe from this list please go to <https://simplelists.susmangodfrey.com/confirm/?u=EEw2b6zZp4y2ifqPiEPjBn8qLVTPRkrG>